Cellular interdependence and collective aspects of the epithelial phenotype : a quantitative and geometric analysis using optical gene activation

Perrine Miquel

THÈSE DE DOCTORAT DE
L'UNIVERSITÉ PARIS DESCARTES

Spécialité
Biologie Interdisciplinaire
École Doctorale Frontières du Vivant

Cellular interdependence and collective aspects of the epithelial phenotype:

A quantitative and geometric analysis using optical gene activation

Présentée et soutenue publiquement par

Perrine Miquel

pour obtenir le grade de
DOCTEUR de L'UNIVERSITÉ PARIS DESCARTES

Soutenance le 16 Novembre 2016, devant un jury composé de :

François Schweisguth	rapporteur
Jean-Loup Duband	rapporteur
Cécile Gauthier-Rouvière	examinatrice
Vincent Hakim	examinateur
Alain Puisieux	exminateur
François Amblard	directeur de thèse
Stéphane Ansieau	membre invité

To my husband,

Abstract

A central idea in biology is that cells are the fundamental units of life, and understanding how the fate and behavior of a cell is established constitute a major task in cell biology. In fact, each cell is under the control of its own determinants (genome, transcriptome, proteome...), but also subjected to some degree to the influence of its neighbors and more generally to its *milieu*. These so-called internal and external determinants both contribute to a cell's fate, leading to possibly highly coordinated cellular communities governed by underlying social patterns. Depending on the relative contribution of internal vs. external determinants, emergent properties are expected to arise. A biological example is the epithelium: the epithelial phenotype itself is not a single-cell phenotype, but emerges from the interaction between potentially epithelial cells.

Our project aims at quantitatively assessing the extent to which the epithelial phenotype relies on internal vs. external determinants (individual vs. social control), by measuring the ability of an epithelium as a cell community to collectively resist the local and heterogeneous induction of the epithelial to mesenchymal transition (EMT). More precisely, we study how the rather well known cause-effect relationship between EMT induction and its transcription signature depends, or not, on the relative position and number of the cells that have been induced and those that have not.

We have established a model where EMT can be optically photoactivated by activating Snail1 part of a fusion protein with the estrogen receptor moiety ERT2 in MDCK cells. Using various geometrical patterns, a neighboring algorithm, cell sorting and qPCR, we show that photoactivation drives a *bona fide* EMT-like effect when homogeneously triggered. However, heterogeneous EMT induction clearly demonstrates that non-induced cells can alter the response of induced cells and this effect can be reversed by varying population densities. The effect of population size and geometry suggests that the next important question is to further characterize the spatial patterning and strength of these social interactions. Understanding underlying rules of epithelial stability open new perspectives to study the onset of carcinogenesis.

Acknowledgment

This work would not have been possible without the support of my thesis supervisor Dr. François Amblard, the help of my colleagues and the understanding of my family and friends.

I am extremely grateful for the quality of the training I have received at the Curie Institute. Joining the team of François Amblard was a great opportunity for me to work in an interdisciplinary team where I deepened my knowledge of cancer biology and broadened my understanding of physics. I would like to thank François for trusting me in taking on this ambitious project and giving me the freedom to explore on my own. The passion for science that drives François was a great source of motivation for me throughout these three years of PhD, and I am thankful for the many insightful discussions that we have had. His absence in the last year and a half of this thesis made the route to the finish line bumpy and tiresome but I terminate this work with the satisfaction that the project will continue on.

I would like to thank the members of my thesis advisory committee, Stéphane Ansieau, Jean-René Huynh and Marco Pontoglio for their advice and ideas along these three years.

In addition, I would like to thank the members of my thesis jury: Jean-Loup Duband, François Schweisguth, Cécile Gauthier-Rouvière, Alain Puisieux and Vincent Hakim. I look forward to an insightful discussion.

L'Institut National du Cancer (INCa) provided the funding for this doctoral work, and gave me the opportunity to explore the fascinating field of cancer. I thank them for their trust and interest in the project.

My progress in the laboratory would not have been possible without the help of experienced colleagues: Thomas Walter for a great collaboration and very enjoyable working sessions, Fahima Faqir for her knowledge in molecular cloning, Zofia Maciorowski, Sophie Grondin and Annick Viguier at the cytometry facility for their patience and expertise, Olivier Renaud and Tristan Piolot for helping me setup the automated patterned illumination experiment.

Thank you to my many friends in the laboratory for making these tough times a positive and pleasant experience: Mélanie, Momo, John, Fahima, Marine, Simon, Sarah, Tatiana, Majdouline...and many more!

I would like to extend my gratitude to my everlasting supportive husband, as well as my parents, brothers, sister and to my friends for their support and encouragements.

And finally, I would like to thank the scientific coordinator of the FdV PhD program, Sofie Leon, for her professionalism, help and kindness in helping me fulfill the many requirements of this thesis. Thank you to the administrators of the laboratory Agnès Verin, Christelle Brochet, Laurence Turpin and Karen Bremond.

List of Figures

List of Table

List of commonly used abbreviations

4-OHT, 4 hydroxy tamoxifen

AJ, Adherens junctions

bHLH, homodimeric and heterodimeric basic helix-loop-helix

BMP, bone morphogenetic protein

BSA, bovine serum albumin

CSC, cancer stem cells

CT, critical threshold

CTC, circulating tumor cells

DTC, disseminated tumor cells

ECM, extracellular matrix

EMT, Epithelial mesenchymal transition

ERT2, estrogen receptor type 2

FRAP, Fluorescence Recovery After Photobleaching

G418, Geneticin

GDF, growth and differentiated factors

GSK-3β, Glycogen synthase kinase 3 β

HIF1α, hypoxia inducible factor 1α

MDCK, Madin-Darby canine kidney

MET, Mesenchymal-epithelial transition

MOI, multiplicity of infection

pEMT, partial EMT

RT, reverse transcription

TGF- βRI/II, type I and type II serine/threonine kinase receptors

TGF-β, Transcription growth factor β

TJ, Tight junctions

ZO-1, Zona occluding 1

ZOD, Zero order diffraction

Contents

Chapter 1

I. Introduction

1 Context of the study

1.1 Interacting elements

Interactions between elements exist in all systems from the macroscopic to the microscopic scales: planets attract each other through gravitational fields, humans evolve in societies, micro-organisms are subjected to cooperation etc... This even holds true for unicellular organisms which are interacting with their environment. Interactions are intrinsically embedded in the universe, and as a result, their existence is often not acknowledged.

In biology, the notion that cells interact and mutually influence each other within a tissue is one of the most fundamental properties of multicellular organisms. During embryonic development, cells are controlled by internal genetic and biochemical determinants, but cues received from neighboring cells are essential for orchestrating developmental patterns [1]. This mutual transfer of information between cells is indispensable and absolutely required for multicellular life. The notion that the fate of cells are under collective control is a key concept which extends beyond the scale of cellular neighborhood. Indeed, the physiology of most organs and systems relies on sets of interacting cellular phenotypes, i.e. cell populations that mutually influence each other and collectively build tissues [2].

1.2 A biological hiatus

In 1839, German botanist Theodore Schwann, came to the conclusion that not only plant, but also animal tissue are composed of cells. He proposed the *Cell Theory* which recognizes the cell as the basic unit of structure and organization in life [3]. This new vision came as a huge breakthrough and led to great development in biology. However, this powerful view presents some limitations when trying to conciliate the intrinsic role of cell interactions with the biology of individual cells. Modern cell biology, biochemistry and genetics, has mostly focused on discovering intracellular structures and determinants controlling the behavior of the cell. Practically however, the knowledge of these determinants is most often inferred from experiments in which cells are not isolated, but rather part of a multicellular ensemble [4] (Figure 1.1, b). While some observables can be individually measured for different cells, e.g. by imaging techniques, many others cannot, and observables are often assessed as population average. In other words, what is often being observed is not how the behavior of an individual cell is controlled by its internal determinants, but rather some relation between the behavior of an "average cell" and the molecular determinants averaged over the cell population (Figure 1.1, c). The relationship between average determinants and average behaviors cannot be used to directly infer the cause-effect relationship at the level of the individual cell. In fact, making this inference relies on the assumption that cells behave linearly, in the sense that cell-cell interactions have no effect on individual cell behaviors (Figure 1.1, c). Because the latter assumption is obviously wrong in general [5], a fundamental issue is raised: to what extent would the knowledge in cell biology hold if experiments had been performed on individual cells, small subsets of cells, or in strongly heterogeneous populations?

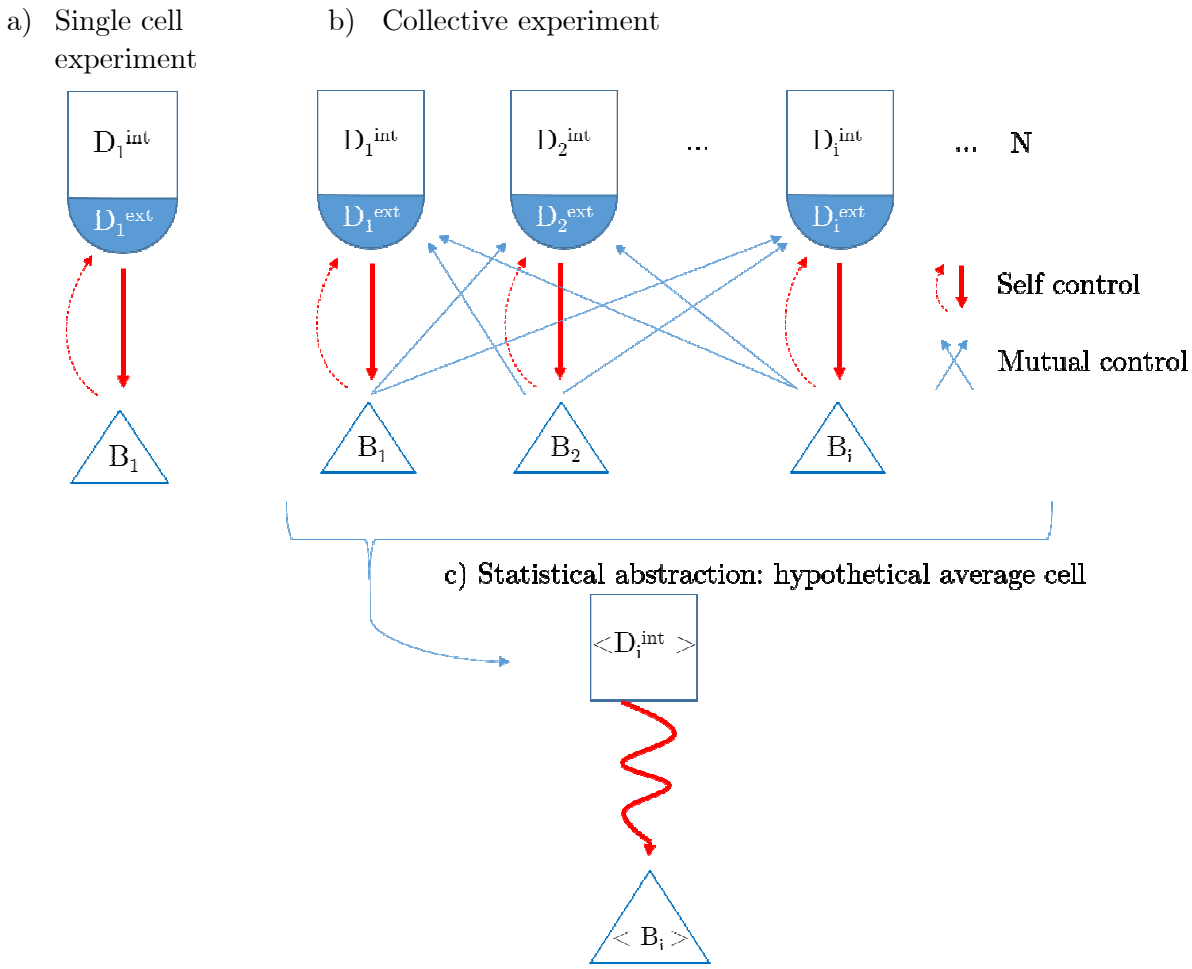


Figure 1.1 Schematic representation of the biological hiatus

Understanding how a particular behavior (B) arises from a set of given internal (D^{int}) and external (D^{ext}) determinants is best conceived in a single cell experiment where the cause and effect relationship is isolated (a). Such a situation is usually not biologically relevant due to the omnipresence of external interactions and in fact, the single cell reasoning is experimentally done in practice on a collectivity of cells (b). In this case, both self-control and mutual influences are at play in dictating the cause-effect relationship. Averaging these sources of influence blurs the distribution between the contribution of internal vs. external determinants, and leads to an *a priori* invalid interpretation unless if cells are fully autonomous (c). There is no reason to believe that the cause-effect relationship, or response element, between the hypothetical single cell experiment and the collective experiment be comparable (compare shape of red arrows).

1.3 The call for a holistic approach

It appears critical to study the degree of autonomy vs. interdependence within cellular communities, but how to isolate autonomous processes from external cues given their degree of relatedness?

Biological systems have witnessed tremendous changes in the viewpoint used to explore them. This comes from the fact, that like any other system, they are associated with a tremendous complexity which scientists try to simplify. In fact, history shows that cell biology has been oscillating between reductionist approaches and holistic approaches [6]. Reductionism attempts to explain a system through the study of its individual parts whereas holism emphasizes on higher level processes that are "more than" the sum of its constituent parts, and the resulting emergent properties [7]. These apparently dichotomic views in fact provide useful information: biological science is fundamentally dualistic. Both reductionist and holistic approaches are necessary and mutually fruitful to bring out the full potential of biological science in understanding living matter [8].

In cancer biology, the reductionist approach has strongly contributed to the discovery of main onco-genes and tumor suppressor genes. In 1971, President Nixon declared the "war on cancer" and launched huge funding campaigns to find a cure for cancer based on genetics. In these days, cancer was though as a disease of identifiable genes and scientists set out to define the molecular signature of cancer cells [6]. In the following decades, although important signaling pathways and key players were characterized, it became clear that genes are not clear-cut "coded instructions" where cause and effect relationships univocally follow out of the DNA sequence [7]. In addition, in the past decade, growing evidence showing the importance of the microenvironment and tumor heterogeneity in cancer progression suggest that a more holistic approach is needed to take into consideration these non-cell autonomous contributions [2, 5, 9]. To date, no simple method has been proposed to generically assess the degree of autonomy vs. interdependence in a cell community.

2 The epithelial phenotype

A reductionist approach is problematic for the study of the epithelial phenotype. Prior to addressing this question, it is important that the biology of epithelia be described. The following section will give insights on the structure, function and regulation of a normal epithelium.

2.1 The normal epithelium

The human body is made up of four different types of basic tissues: epithelial tissue, connective tissue, muscle tissue and nervous tissue. Each is formed by the assembly of specialized cells that allow for specific structures and functions. The focus of this section will be on epithelial tissues which line the cavities and surfaces of most organs and blood vessels in the body [10].

2.1.1 Structure of the epithelium

Epithelial tissues are made up of cohesive and polarized epithelial cells that can be of different kind: squamous, columnar and cuboidal. These differentially shaped cells can assemble to form either simple single-layered, stratified or pseudostratified epithelia [11] (Figure 2.1). Epithelia are characterized by their strong cellular cohesiveness, their ability to play a barrier-like role and a well-established apico-basal polarity. Epithelial junctional complexes are symmetrical structures formed by adjacent cells that allow for these specified roles and consist of a number of adhesion types.

Simple epithelia can be thought of as a 2D structure in which all cells play the same structural role. More complex situations are found with stratified epithelia in which the successive cell layers are not equivalent. The very notion of polarity, in this case, is different at the level of the individual cells, but the multilayer structure still exhibits a functional polarity often associated with a differentiated gradient.

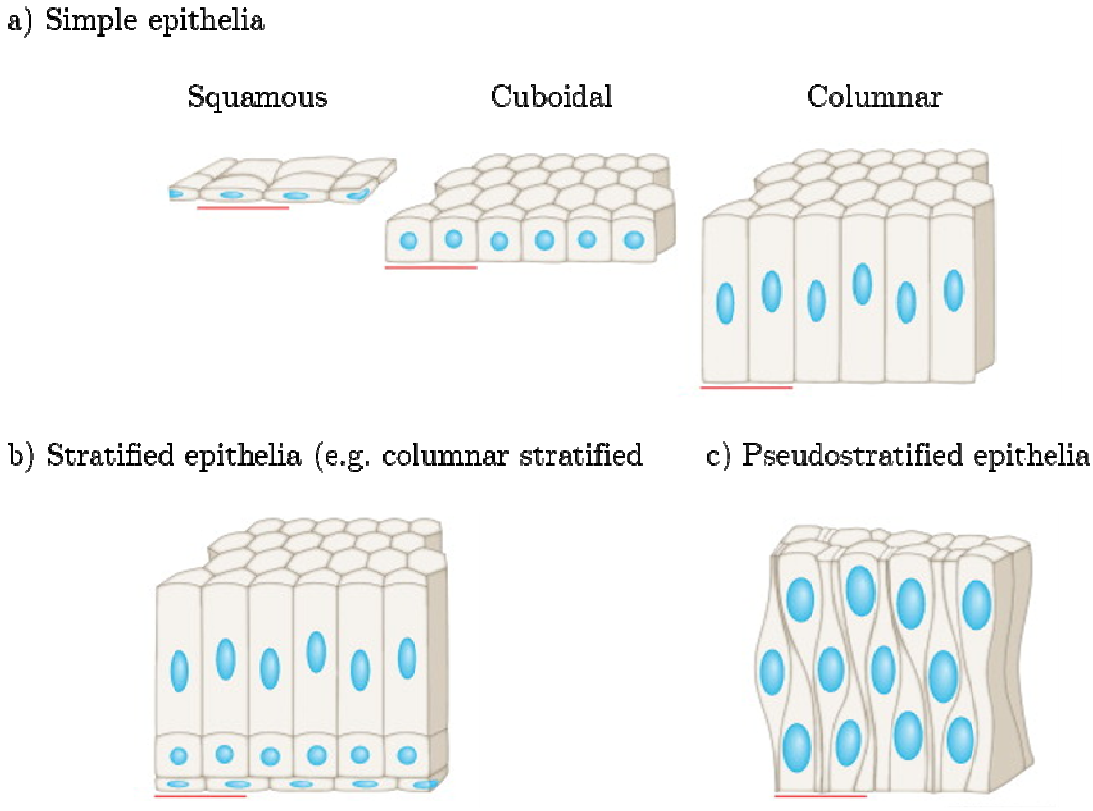


Figure 2.1 Representation of the different types of epithelia

Simple epithelia are composed of a single layer of cells (a) whereas stratified epithelia consists of several layers one on top of the other (b). In both cases, the epithelia is defined by the shape of the cells composing it. Squamous cells are thin and relatively flat, cuboidal cells have a cubic shape, and columnar cells are thin and tall. Pseudostratified epithelia are composed of a single layer of cell but the nuclei of these cells are positioned in a manner suggestive of stratified epithelia. From [11].

In the present dissertation and work, only the 2D assembly of simple epithelia is considered.

2.1.2 Intercellular junctions

Adherens junctions (AJ) play a fundamental role in maintaining tissue organization by allowing cell-cell anchoring to one another as well as to extracellular matrix. AJ include focal adhesion, desmosomes and hemidesmosomes junctions. The main proteins involved in AJ are catenins and cadherins and more specifically, E(epithelial)-cadherin (encoded by the *CHD1* gene), N(neural)-cadherin and P(placental)-cadherin [12]. These single membrane spanning proteins have a conserved cytoplasmic domain and a divergent extracellular domain of five repeats. By interacting with cytoplasmic catenin, cadherins promote calcium dependent homophilic cell-cell adhesion. More specifically, β-catenin interacts with the cytoplasmic region of the

cadherin. α-catenin then binds to β-catenin and links the cadherin to the actin microfilaments of the cytoskeleton [12]. This indirect coupling allows for strong cell-cell adhesion as lateral clustering of cadherin complexes develop [13]. The coupling of AJ with the cytoskeletal machinery of the cell highlights that cadherins are not simply "biological glues" but rather can mediate signals from the cell surface to influence processes such as cell shape, cell polarity, apoptosis or cell migration [14, 15]. A sub-family of AJ are desmosomes, known as spot-like dynamic structures that resist mechanical stress and are engaged in calcium independent adhesion [16, 17]. Due to their hyper-adhesive nature, they are very abundant in epidermis and myocardium [18]. Desmosomes may regulate, through desmosomal cadherin, the availability of signaling molecules thus taking part in fundamental processes such as cell proliferation, differentiation and morphogenesis. Similar in form, hemidesmosomes allow to bind the basal portion of an epithelial cell to the underlying basement membrane, a specialized form of extracellular membrane (ECM), through integrin recruitment. Taken together, AJ play an important role in establishing tissue integrity as well as acting as acceptors and effectors of cell signaling pathways [19] (Figure 2.2).

Tight junctions (TJ) are a second class of structures that allow for epithelial organization. TJ are essential to form apical semipermeable diffusion gates for ions and solutes between cells. Located apically, they allow for selective diffusion based on charge and size and are essential for the maintenance of tissues and organs homeostasis. In addition, TJ ensure a barrier function to prevent the mixing of apical and basal membrane proteins [20]. The complex architecture and composition of TJ has rendered their molecular and functional understanding quite challenging [21] but proteins taking part in the establishment of TJ have been characterized. In fact, the main players in these barriers are transmembrane proteins that are part of the claudin and occludin family [22]. Claudins are composed of four transmembrane helices which include a variable cytosolic tail. The carboxylic terminal is characterized by a typical PDZ-binding motif which have been shown to bind to PDZ domains of the cytoplasmic scaffolding proteins ZO-1, -2 and -3 [23]. The latter are known to bind directly to actin filaments through their COOH terminal regions, thus acting as a cross linker between TJ and the cell's cytoskeleton [23] (Figure 2.2).

Finally, gap junctions form a third type of adhesion structures. These specialized intercellular connections link the cytoplasm of two cells, allowing for the direct passage of ions and small molecules from one cell to the other. This organization requires that two hemichannels made up of connexin proteins from adjacent cells come together to form a functional central pore. This association can be homo or heterotypic [24]. Gap junctions

are key in the ability of adjacent cells to coordinate electrical and mechanical outputs [25] (Figure 2.2).

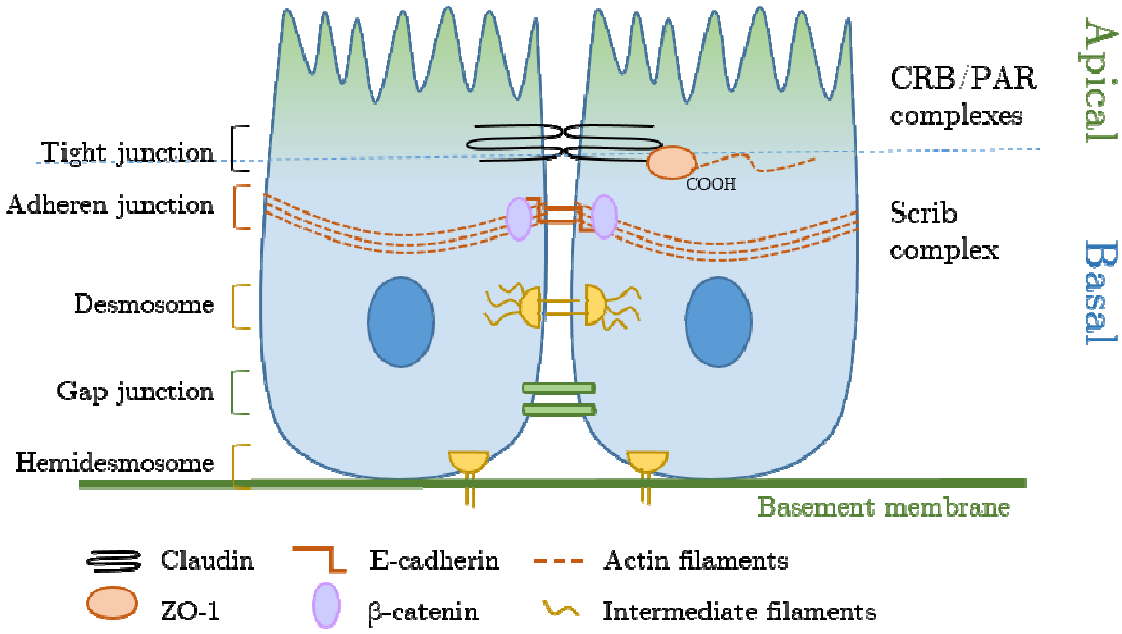


Figure 2.2 Adhesion complexes in epithelial cells

Schematic representation of the different adhesion complexes involved in epithelial intercellular junctions. Tight junctions define the limit between apical and basal compartments in the cell and are formed by Claudin proteins. Adherens junctions are formed mainly by E-cadherin which are maintained at the membrane by β–catenin. The latter act as a cross linker with the cell's cytoskeleton. Desmosomes and hemidesmosomes are hyper adhesive structure that are composed of integrins. Finally gap junctions form direct passages for solutes between two adjacent cells. Adapted from [26]

2.1.3 Polarity

Beyond lateral organization, epithelial cells also have a well-established apico-basal polarity axis. This is critical for cells to distinguish the interior of the organism from the luminal or external environment. In fact, epithelial cells have distinct molecular components along this vertical axis: an apical domain facing the lumen and a basolateral region where adhesion complexes hold adjacent cells together as well as to the underlying basement membrane [27]. In fact, TJ, which locally sequesters proteins, contribute to defining the apical compartment although their role in polarity is still unclear [28]. Molecularly, three main evolutionary conserved groups of proteins are known to play a major role in the establishment and maintenance of epithelial cell polarity. The Crumbs (CRB) complex, the PAR system and the Scribble (Scrib) module are segregated along

the apico-basal axis and locally contribute to this asymmetrical pattern of organization [29]. Genetic studies on *Drosophila* have shown that mutants for Scrib module show opposite phenotypes to those observed for mutants of CRB or PAR complexes. This suggests that Scrib and CRB/PAR complexes have antagonistic activities [30] and that they are essential to setting up apical and lateral/basal identities [28]. The exact mechanism of action of these complexes and their cross regulation are not yet fully understood, but it has been speculated that the actin cytoskeleton of the cell may play a role [28, 31] (Figure 2.2). This structural polarity allows the localization of specialized proteins, leading to a functional polarity: various ion pumps and transporters are found apically and allow for the functionalization of the epithelium and its vectorial transport capacities [1].

2.1.4 Function of the epithelium

Epithelial cells with apico-basal polarity and established intercellular adhesion complexes form dynamic barriers to protect the organism from the external environment. Besides compartmentalization properties, epithelial cells are also known to have a secretory potential. Certain types of epithelia include cells that exhibit more diverse functions. For instance, the gut epithelium is endowed with invaginated folded structures that are associated with endocrine/exocrine cells. These cells secrete enzymes, hormones, and surfactants which are delivered to the rest of the epithelium or apical fluid compartment [32]. Epithelial cells that make up the lining of the gut are important for the absorption of nutrients into the blood stream. In addition, mechanical fluctuations are exerted by some epithelia via apical cilium. These actin-rich structures are key in sensing the external environment through chemo-, thermo-, or mechano-sensation and can mechanically create a fluid flow in the apical compartment. Typical ciliated epithelial cells can be found in airway epithelium where they are key in clearing respiratory mucus [33]. In general terms, epithelial integrity, and intercellular adhesion are key for these tissue to exert their primary functions.

2.1.5 The epithelial particularity

This broad description of the epithelium highlights the degree of organization and the various structural and functional features of most epithelia of many organisms. Interestingly, none of the characteristic properties of the epithelium are observed when an epithelial cell is isolated. The key requirement for the existence of the epithelial phenotype is the formation of intercellular junctions with neighboring cells: it is not possible for an epithelial to polarize itself in absence of junctions. As a result, an individual epithelial cell will not be able to carry out its normal functions. This is not as true for the other main types of cells. In fact, a single muscle cell retains its main

property regardless of the presence or absence of neighbors: individual cardiomyocytes can beat [34]. All the same, although individual neurons cannot form synapses, the architecture of the cell is intact and the axon-dendrite polarity is established [35]. These observations point to the fact that, unlike other cell types, the epithelial phenotype is uniquely a non-autonomous process.

Most specifically, in the experimental situation where an epithelial cell line is seeded at very low density in culture in such a way that cells are initially well separated, no junctions can initially be made, nor can polarization be established. This "lonely epithelial cell" is in a situation of self-control, and so is not *actually* epithelial. In fact, distinguishing this cell from a dispersed mesenchymal cell would be very difficult. However, as cell density reaches the point where junctions can be made, the lonely epithelial cell will engage in epithelial junctions that other non-epithelial cells are not capable of establishing. Isolated epithelial cells are *potentially*, but not *actually* epithelial, leading to the notion that they contain distinctive internal epithelial determinants (genetic, biochemical). At confluency, these internal determinants are combined with external determinants (contact signals, secreted signals?) in a congruent fashion leading to the full blown epithelial phenotype.

To fully understand the epithelial phenotype, it is thus needed to elucidate the crosstalk of external determinants with the internal determinants. A collective approach is required to resolve the relative contribution of these two sources of information. In the congruent situation, it will be difficult to resolve these contribution from one another, but this task should be easier if the situation is no longer congruent. Indeed, the outcome of the competition between antagonistic causes should help to determine which of the two dominates. This reasoning led to investigate the individual vs. collective nature of the epithelial phenotype by challenging the epithelial stability using the epithelial mesenchymal transition (EMT).

2.2 Epithelial to Mesenchymal transition

The epithelial-mesenchymal transition (EMT) refers to a complex molecular program whereby epithelial cells progressively lose their characteristic phenotype to become mesenchymal cells. The first observation of EMT arose in the early 1980s, when Elizabeth Hay, pioneer in the field of EMT, described this phenotypical switch in the primitive streak of the chick embryo [36, 37]. The process of EMT has been proven to be essential for germ layer formation and cell migration in the early vertebrate embryo [38].

The cascade of events associated with EMT are critical in normal embryonic development and in adult wound healing, tissue regeneration and fibrosis [39]. In

addition, as will be detailed in further sections, the EMT machinery has been shown to be aberrantly reactivated in certain cases of cancers, contributing to the progression of the pathology.

2.2.1 EMT: a cascade of events

The process of EMT can be described in several key events taking place once epithelial cells become responsive to EMT-inducing signals. This involves disruption of cell-cell adhesion complexes; loss of apico-basal polarity leading to rear-front polarity; cytoskeletal reorganization and changes in cell shape; increased cell protrusion and motility; ability to degrade the extracellular matrix (ECM) to favor invasion. Molecularly, EMT is associated with a switch in the expression of epithelial genes part of adhesion complexes (such as E-cadherin, claudins and occludins) to the expression of mesenchymal genes (such as N-cadherin, fibronectin, vimentin) [40] (Figure 2.3).

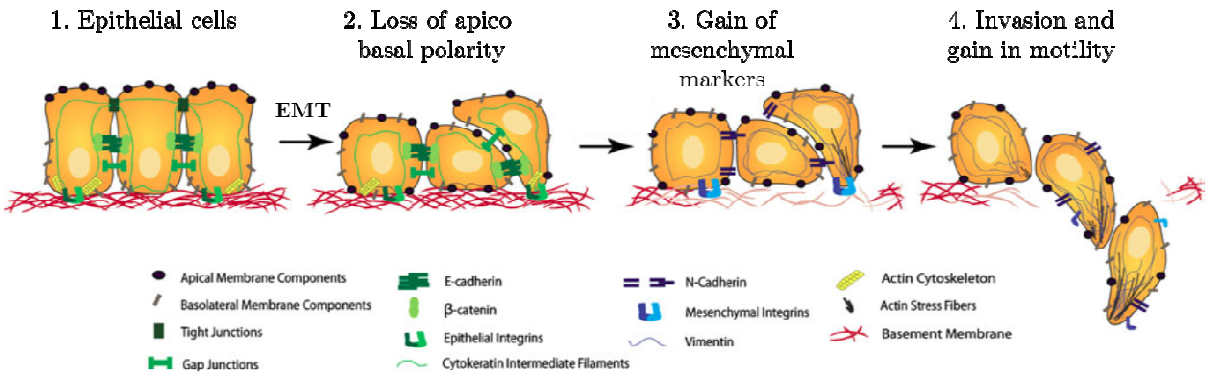


Figure 2.3 Epithelial Mesenchymal Transition

Epithelial cells undergo drastic morphological remodeling during EMT. The loss of apico-basal polarity result from adhesion complex disassembly (2). E-cadherin, cytokeratins and epithelial integrins (green) are downregulated and replaced by mesenchymal proteins such as N-cadherin, mesenchymal integrin and vimentin (blue) (3). Actin cytoskeleton remodeling lead to accumulation of stress fibers at membrane protrusion sites. Ultimately, the underlying basement membrane is degraded forming a breach that allows cells to individually move into the surrounding stroma (4). Adapted from [40].

EMT occurs in distinct biological settings that lead to very different functional outcomes. This has led to the classification of EMT in three classes: type 1 EMT relate to events engaged in embryogenesis and normal organ development, type 2 EMT refer to wound healing and tissue repair, and type 3 EMT is associated with pathological cancer progression and the formation of migratory metastatic cells [41, 42]. Although these three

classes of EMT are described as distinct biological events, the underlying molecular players involved are shared between these outwardly diverse phenotypic programs [43].

2.2.2 EMT in normal development

The first instances where EMT occur during development is in gastrulation. Initially, the epiblast forms a regular epithelium lying on a basal lamina. Gastrulation is required for mesoderm formation at a precise region of the primitive ectoderm referred to as "primitive streak" which will mark the posterior and anterior portions of the embryo. Under the control of specific signals produced by the Spemann-Mangold organizer, the cells part of the primitive streak will lose their epithelial organization and undergo mesenchymal differentiation. These now motile cells are ready to ingress, and migrate away in the space below the ectoderm through the locally breached basal lamina [40, 43, 44] (Figure 2.4, a).

Later in development, a second example of EMT arises in the generation of the neural crest. During closure of the neural tube, cells located at the border of the embryonic ectoderm and the neuroectoderm undergo EMT, thus generating a migratory neural crest cell population [45]. These cells can then invade from their site of origin into the developing embryo to give rise to diverse lineages including melanocytes, craniofacial cartilage and bone, smooth muscle, peripheral and enteric neurons and glia [40, 43] (Figure 2.4, b).

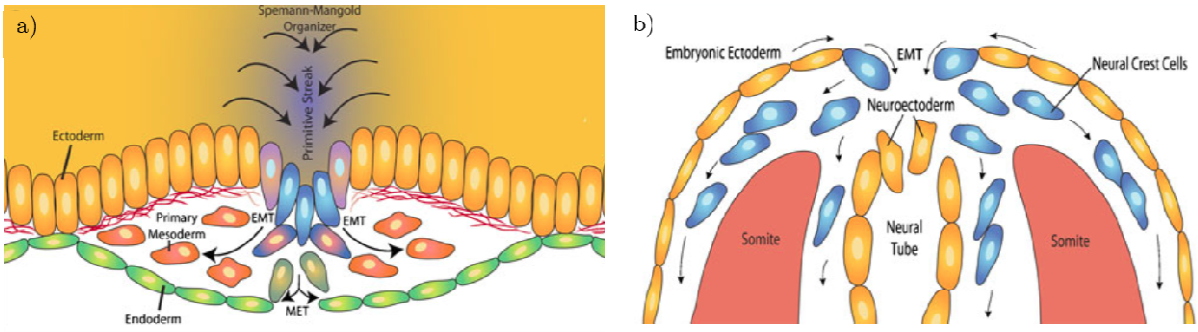


Figure 2.4 Developmental EMT

In normal embryonic development, epithelial cells undergo EMT. Here are depicted two very well characterized examples. a) Representation of mesoderm formation during gastrulation. Epithelial cells of the primitive streak undergo EMT and ingress into the underlying tissue. The migratory cells either give rise to the primary mesoderm or undergo mesenchymal-epithelial transition (MET) to form the endoderm. b) Closure of the neural tube induces cells at the border of the neural plate and of the ectoderm to undergo EMT. These migratory cells will invade the developing embryo and give rise to various differentiated cells. From [40].

The phenotypic transition associated with EMT is not an irreversible commitment. In fact, a similar but reverse process known as mesenchymal-epithelial transition (MET) allows dynamic cellular remodeling [41]. Cycling of EMT-MET is instrumental in kidney ontogenesis. The earliest structures making up the kidney are the uretic bud and the nephrogenic mesenchyme which form the collecting duct and nephron respectively. During kidney development, the nephrogenic mesenchyme form aggregates around the uretal bud and begin to express laminin and epithelial markers. In addition, the cells become polarized and differentiate into epithelial cells [40, 46].

Overall, primary EMT such as gastrulation and neural crest formation allows for the generation of motile cells which can then undergo EMT to form secondary epithelia cells that support organogenesis [42]. Structures derived from the mesoderm are later remodeled by successive EMT-MET cycles which will define the architecture of different organs and tissues. It is important to highlight the notion that EMT and MET refer to phenotypic transitions that affect cell shape and adhesion property rather than specification or acquisition of cell fate itself [43].

2.2.3 Inducers of EMT

The most potent inducers of EMT are members of the transforming growth factor β (TGF-β)/bone morphogenetic protein (BMP)/growth and differentiated factors (GDFs) family. More specifically, TGF-β isoforms (TGF-βs) play an important role during embryogenesis and tissue homeostasis and are considered to be prototypical cytokine for the induction of EMT [47]. TGF-β exert its effect through binding to type I and type II serine/threonine kinase receptors (TGF- βRI/II) [48]. Upon TGF-β stimulation, the receptors are internalized into early endosomes and TGF-βRI phosphorylates intracellular signaling effectors Smad2 or Smad3. The latter then form a complex with Smad4 and translocate into the nucleus where, in conjunction with other EMT transcription factors (EMT-TF), they regulate specific genes [47, 48]. In fact, TGF- β induced Smad activation causes direct activation of the mesenchymal genes fibronectin, vimentin and collagen α1 (Figure 2.5) [26]. Parallel to this Smad-dependent signaling, TGF- β also mediates its effects through Rho-like GTPases, PI3K and MAPK pathways [49]. Phosphorylation of the adaptor protein Par6 by TGF- βRII causes direct post translational modification and regulation of proteins involved in intercellular junctions [26, 50] and drive actin reorganization and lamellipodia/filopodia formation.

EMT can also be induced by the Wnt and Notch pathways as well as with other growth factors such as fibroblast growth factor (FGF), insulin growth factor (IGF) or epidermal growth factor (EGF) [26]. Suffice to say that different extracellular factors often activate several EMT pathways which regulate the expression of common targets of

EMT. This suggests the existence of signaling cooperation and pathway crosstalk commonly regulating downstream EMT effectors [26] (Figure 2.5).

2.2.4 EMT-TFs

Repression of the epithelial phenotype and activation of the mesenchymal phenotype in EMT result from changes in the expression levels of master regulators known as EMT transcription factors, or EMT-TFs. Their early activation during EMT confers them a central role for all three classes of EMT. Many of the target genes of these EMT-TFs are common, allowing for a coordinated repression of epithelial genes. These include the Snail zinc family, the zinc finger E-box family, and the homodimeric and heterodimeric basic helix-loop-helix (bHLH) family, many of which are frequently overexpressed in certain cancers [26, 51]. Loss of functional E-cadherin is a hallmark of EMT and these EMT-TFs are key in triggering dynamic silencing of the *CDH1* gene [52].

Snail family

The Snail family which include Snail1 and Snail2 (also known as Slug) in vertebrates, have been shown to be associated with all known EMT events [42]. Snail proteins are potent repressors of epithelial genes, and act through binding of their carboxy-terminal zinc finger domains to E-box DNA sequences of target genes. E-box sequences are formed by the consensus palindromic sequence CANNTG [52]. More specifically, Snail1 which is activated by TGF-β, Wnt, Notch and growth factors pathways, induces direct repression of E-cadherin by binding to the proximal promoter region and inducing histone modifications [53]. Snail1 activity is regulated through post translational modifications which dictate the localization, and thus activation, of the protein [26, 54]: Snail1 must translocate to the nucleus to be active [55]. Glycogen synthase kinase 3β (GSK-3β) can phosphorylate Snail1 in two distinct serine rich motifs. Phosphorylation causes inactivation of the transcription activity of Snail1 through nuclear export and proteolytic degradation signals [56]. Modulating GSK-3β mediated phosphorylation is used by several pathways to increase Snail1 activity [26]. This EMT-TF is a highly unstable protein with as half-life of 20 to 45 minutes that undergoes rapid turnover [57]. Tight regulation through spatial localization dictates the ability of Snail1 to repress epithelial genes. Aberrant nuclear transport mechanism are commonly found in cancer and can shift this fine-tuned compartmentalization leading to abnormal Snail1 activation [55].

Zinc finger E-box family

The zinc finger E-box family is composed of two members in vertebrate: ZEB1 (also known as δEF1) and ZEB2 (also known as SIP1). Similarly to the Snail family, ZEB are induced by TGF- β, Wnt and growth factor activated Ras-MAPK signaling pathways

[58]. Snail1 is known to be a direct activator of ZEB1 such that its activation is more downstream than Snail1 in the EMT induction pathway. ZEB members are subjected to post translational modulation through the repressing activity of small non-coding mRNA, or microRNA (miRNAs). In particular, the miR200 and miR205 family are engaged in a double negative feedback loop with ZEB members [59].

Homodimeric and heterodimeric basic helix-loop-helix (bHLH) family

The bHLH family composed of TWIST1, TWIST2, E12 and E47 form the third family of EMT-TF that play a strong role in epithelial gene repression and mesenchymal gene activation. Unlike ZEB1 activation which is downstream of Snail1, TWIST1 can induce direct repression of E-cadherin and activation of fibronectin and N-cadherin in a Snail independent manner [42]. TWIST1/2 bind to DNA as homo- or heterodimers to regulate E-box DNA sequence [57]. TWIST are induced by the various pathways described previously, but most notably by hypoxia inducible factor 1α (HIF1α). This plays a critical role in hypoxic tumor environment thus favoring cell dissemination. TWIST expression has also been shown to be induced by mechanical stress, in a β-catenin dependent, in *Drosophila melanogaster* [60].

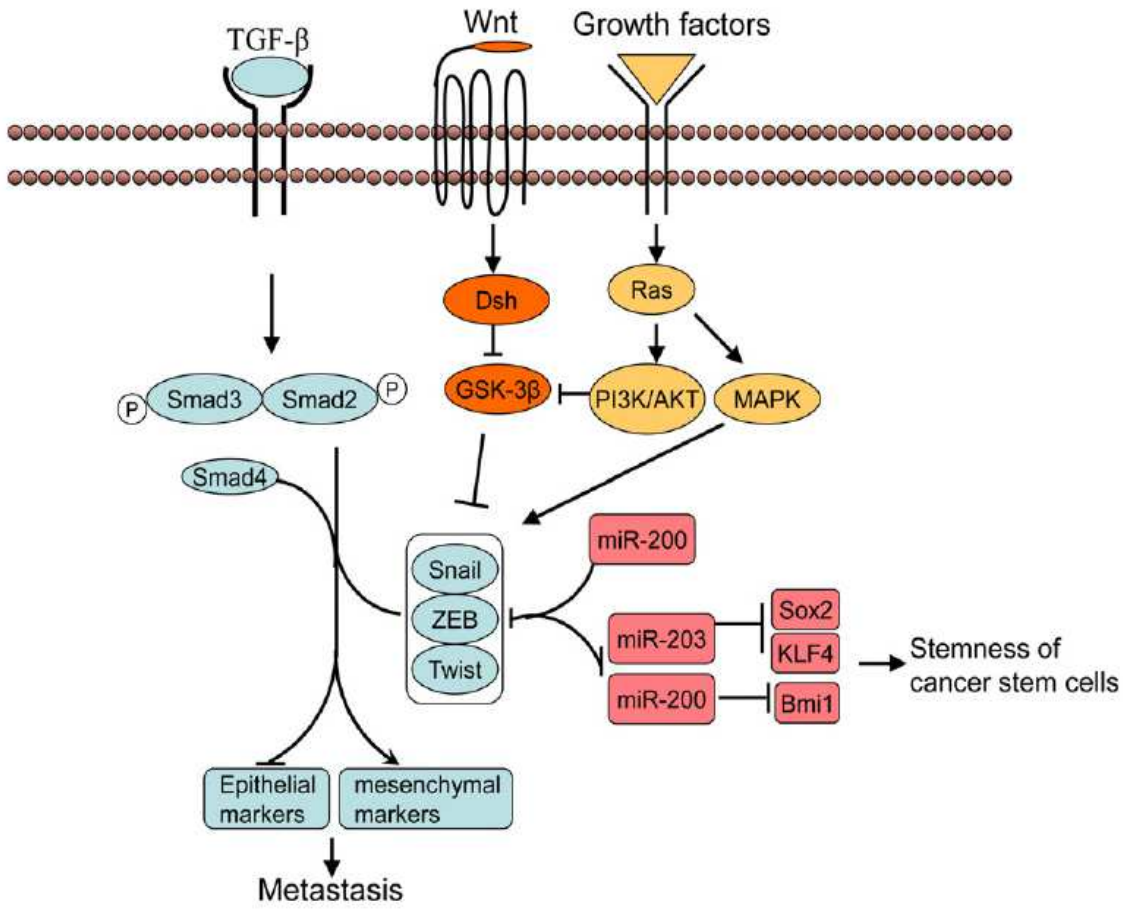


Figure 2.5 Signal transduction pathways associated with EMT

Schematic representation of the different activation pathways of EMT. TGF-β is a major inducer of EMT through phosphorylation and activations of Smad complexes. When translocated into the nucleus, these complexes interact with the three main EMT-TF: Snail, ZEB and Twist to regulate downstream expression of TGF-β target genes. Wnt and growth factors can also induce EMT, acting through receptor tyrosine kinases in pathways that cooperate with the TGF-β signaling. EMT-TF are known to be negatively regulated by miRNAs. From [49].

The expression patterns of Snail, ZEB and bHLH families are relatively similar in their activation, regulation and downstream targets. Complex chemical and mechanical signaling from the microenvironment can coordinate the expression of these factors and promote their interplay in orchestrating E-cadherin repression. Crosstalk loops exists among EMT-TFs and it has been shown that Snail1 plays a more prominent role in initiating EMT while ZEB and Slug are rather involved in maintaining the mesenchymal state [52, 61] (Figure 2.5).

2.2.5 EMT reporters

Detection of an EMT event can be assessed morphologically and through the changes in expression and/or localization of a variety of biomarkers, or effectors. In fact, cytoskeletal reorganization upon EMT causes drastic changes in the expression of many cell surface and cytoskeletal markers. E-cadherin, which has been touched upon in the previous section, has long been considered the hallmark of EMT through a strict correlation between the loss of the epithelial phenotype and the loss of its expression. However, this correlation is highly questionable. It has been suggested that while E-cadherin is necessary for intercellular junction formation, it would be dispensable for the maintenance of mature junctions [62]. Conversely, some reports show that the loss of E-cadherin at the junction is not a necessity for EMT to take place [63, 64]. This point is part of an ongoing debate on the histological relevance of using E-cadherin repression as a main readout of EMT. Suffice to say that the presence of E-cadherin at the junctions does not imply that they are engaged in adherens junctions. The protein could alternatively remain in a juxta-membranous localization that one would hardly distinguish from engaged cadherins. In situations where E-cadherin is actually repressed, there is a switch from E to N cadherin, and the latter can be used as a reporter [65]. A similar integrin switch occurs during EMT, and can be used to assess the alterations of cell-ECM interactions [42]. However, this reporter is strongly cell-dependent and thus has limited utility. As a result of tight junction dissociation, expression of claudins and occludins can be useful to appreciate the advancement of EMT. Cytoskeletal markers include vimentin which is highly expressed in mesenchymal cells and fibroblast. Expression of vimentin reflects the degree of invasiveness of the cells [42]. In addition, α-smooth muscle actin (α-SMA) will reorganize into stress fibers as a result of cellular elongation and increased motility. Finally, β-catenin is released from E-cadherin into the cytosol and into the nucleus when intercellular junctions are lost in EMT [42]. Fibronectin is a high molecular weight protein which serves as a scaffold for fibrotic ECM and whose levels increase during EMT. The use of fibronectin as a reporter of EMT is dependent on the cell type considered [42]. The main reporters of EMT are summarized in (Table 1).

Acquired markers	Attenuated markers
Cell-surface proteins	
N-cadherin	E-cadherin
OB-cadherin	ZO-1
α5β1 integrin	
αVβ6 integrin	
Syndecan-1	
Cytoskeletal markers	
FSP1	Cytokeratin
α-SMA	
Vimentin	
β-Catenin	
ECM proteins	
α1(I) collagen	α1(IV) collagen
α1(III) collagen	Laminin 1
Fibronectin	
Laminin 5	
Transcription factors	
Snail1 (Snail)	
Snail2 (Slug)	
ZEB1	
CBF-A/KAP-1 complex	
Twist	
LEF-1	
Ets-1	
FOXC2	
Goosecoid	
MicroRNAs	
miR10b	mir-200 family
miR-21	

Table 1 Summary of the main reporters of EMT

Non exhaustive list of markers of EMT. Note the striking imbalance in the vast list of markers of the epithelial phenotype with comparison to the sparse number of markers of the mesenchymal state. Adapted from [42].

Finding a single EMT reporter is impossible because the transition is a multidimensional process. Rather, a set of various biomarkers should be considered to appreciate the transition. It is interesting to note the great asymmetry between the broad list of reporters undergoing an increased expression upon EMT, vs. the shy number of repressed reporters.

2.2.6 Models of EMT

Decades of research on EMT have led to various different experimental models both *in vivo* and *in vitro*. The main animal based models to study developmental EMT are derived from *Drosophila*, amphibian and avian embryos [36, 37, 45]. Mice models are commonly used to study pathological EMT through explants and cell injections, and provide a physiological system to address the role of the microenvironment. *In vitro* models are most commonly based on primary and immortalized stable cell lines. The choice of the model to use is based on the biological question asked: different models allow to address different questions. For example, the study of migration associated with EMT is better addressed in *in vivo* models as the artificial substrate in *in vitro* model is not adequate to conclude. *In vitro* models do not recapitulate the whole biological event of EMT, but allows to study EMT as a dispersion, or loss of polarity phenomenon.

The main interest of considering an *in vitro* EMT model in this thesis work is that it offers a mean to create a situation of non-congruency using heterogeneous induction of a physiological phenotypical transition.

2.3 Emergent properties of the epithelium

2.3.1 Mechanical stability of the epithelium

From the mechanical point of view, the junction between two cells is a rather complex object. Indeed, in recent work [66], it was shown that adhesion structures are very dynamic due to E-cadherin endocytosis. The life time of E-cadherin residence at the adherens junction is in fact under the control of endocytosis, at a rate that is controlled by the adhesion and contractile forces. Interestingly, this rate varies in time and space, but it is mechanosensitive and is the same across individual junctions. This report strongly suggests that junction represent autonomous and symmetric objects, part of a joint control between two adjacent cells. It is as if mutual control prevails over cell autonomous control.

These observations suggest the possibility to consider the epithelium not only as an ensemble of interacting cells, but alternatively as a lattice of interconnected junctions. Incidentally, a number of processes such as intercalation and convergence extension are

governed by the contractile properties of these interconnected junctions and large scale transcellular contractile structures. This clearly illustrates the dual nature of epithelia: a set of cells with partially autonomous properties, or a lattice of individual junctions [67, 68] (Figure 2.6).

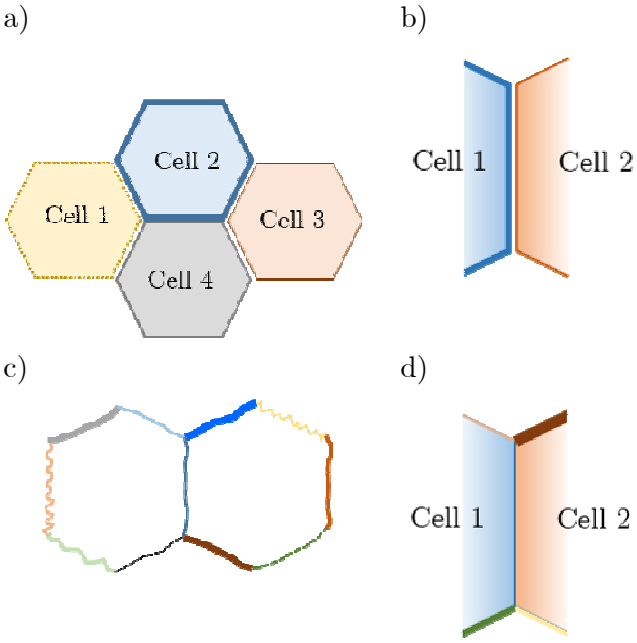


Figure 2.6 Dual nature of the epithelial cell

Epithelial cells are most commonly seen as a community of cohesive autonomous cells, defined by their internal determinants. In this case, junction dynamics is under cell autonomous regulation (a) and junctions are heterogeneous (b). This is problematic when it comes to understanding collective stability. Epithelial cells can thus also be seen as a lattice of junctions which dominate over cellular individuality (c). In this case, a given junction is unique and relies on the ability of the two partner cells to stabilize it (b).

The notion of an epithelium as a lattice of junctions comes with the notion that contractility continuously operates from subcellular to supracellular scales, as if the individuality of the cellular control was unimportant.

2.3.2 EMT, cancer and stemness

Aberrant reactivation of the EMT developmental program has been shown to play a critical role in promoting metastasis in epithelial-derived carcinoma and marks the first step of the "invasion metastasis cascade" [69]. This multistep process starts with the loss of apico-basal polarity and the gain in motility, necessary for the initial escape from

the primary site. Breakdown of the basement membrane initiates invasion and allows for migratory cells to detach from the primary tumour and enter the systemic circulation. In this new environment, only a few cells are able to survive as circulating tumour cells (CTCs), and extravasate through capillary endothelium. Once settled in target tissues, CTC are termed DTC (disseminated tumour cells) and can colonize or undergo MET [70]. This last step can be followed by a re-epithelization *via* a MET, and the ultimate colonization of distant organs and formation of secondary tumours [41, 71]. Pathological EMT is very similar to physiological EMT in terms of signalling pathways, regulators and effector molecules [72] (Figure 2.7). EMT/MET enable solid tumours, over 90% of which are epithelial in nature [73], to disseminate and colonize distant organs. The cellular hijacking of the normal developmental process for metastasis is the cause of 9 out of 10 cancer related deaths [41, 74].

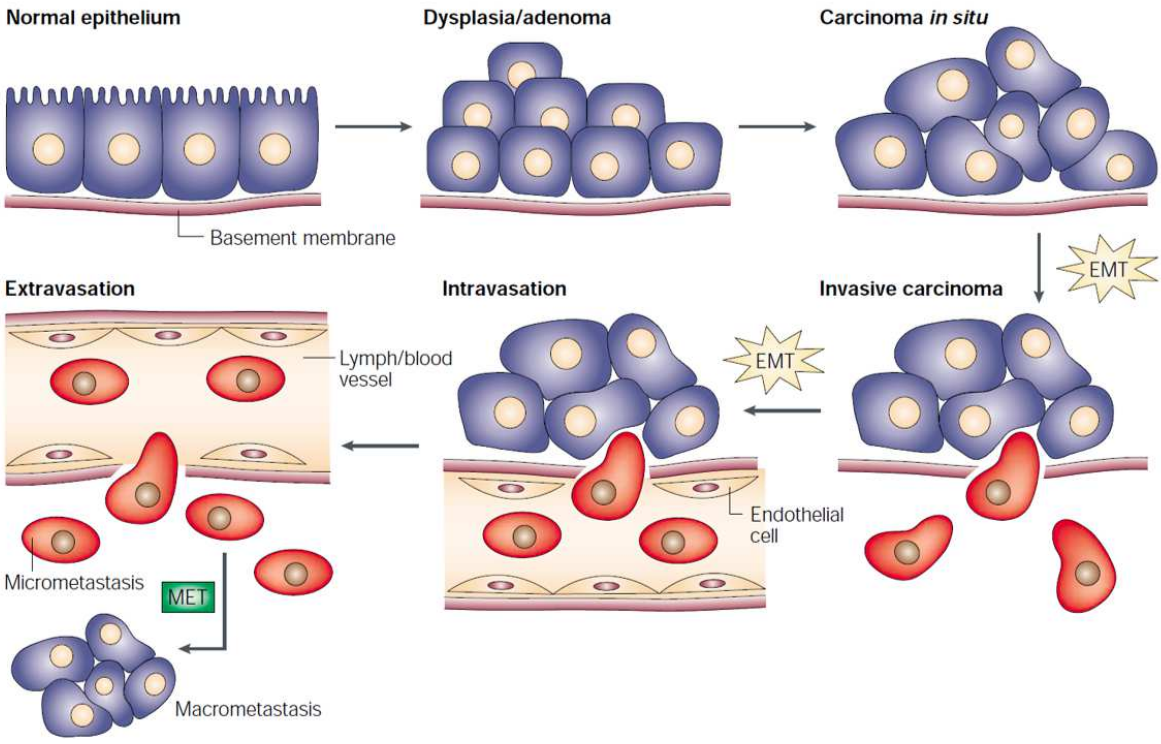


Figure 2.7 EMT and MET in the emergence and progression of carcinoma

Epithelial cells can become dysplasic as a result of epigenetic and genetic alterations. Progression of the dysplasia can lead to *in situ* carcinoma. EMT induces basement membrane degradation, allowing intravasation of motile cells to reach the blood circulation. Solitary carcinoma cells can extravasate from the blood circulation to colonize secondary sites through MET (From [69]).

Importantly, EMT and MET, whether in physiological or pathological contexts, are not binary or unidimensional processes [75]. In fact, CTCs have often been observed as clusters of circulating cells with hybrid epithelial (cell-cell adhesion) and mesenchymal (motility) traits [75, 76], suggesting that EMT is not an all or none event but rather a gradient which can lead to partial EMT (pEMT) [77, 78]. The hybrid phenotype is associated with collective cell migration, efficient exit from the bloodstream, resistance to apoptosis and render the cells much more metastatic than individually migrating CTCs that have a complete EMT phenotype [75].

Stemness is the ability of a cell to proliferate in an asymmetric way allowing for the maintenance of an undifferentiated reservoir as well as the generation of new differentiated cells [79]. Cells with self-renewal properties are referred to as cancer stem cells (CSCs) and were first identified in the hematopoietic system [80]. Induction of EMT in human mammary epithelial cell line has been shown to be associated with the expression of stemness markers such as $CD44^{high}/CD24^{low}$ [81]. Extensive literature now supports the idea of a link between the EMT program and the stemness phenotype. The existence of CSCs explain, in part, how disseminated cells, or CTCs, are able to proliferate and form macroscopic metastasis in distant organs [81]. CSCs are thought to play an important role in drug resistance as they naturally express drug efflux pumps and as such, persist in tumors and in metastasis, contributing to relapse [82].

Cells with a hybrid phenotype are pointed out as the "bad actors" of cancer metastasis as their role in drug resistance, stemness and cancer relapse is being revealed [75]. Understanding cell-fate decision of these hybrid intermediates engaged in partial EMT is crucial for the anticipation of cancer progression and prognosis outcome. At the stage of DTCs, if one considers that these decisions are to some extent non-cell autonomous, their molecular determinants are likely to be not only internal but also external. A major issue is to know which of internal vs. external determinants dominate.

2.3.3 EMT, a model for the study of emerging properties

The description of the epithelium and the blurred limit between cellular individuality and collective epithelial phenotype points to the fact that EMT cannot be a single cell process and constitutes a great experimental framework to measure the degree of cellular interdependence: the response to EMT *cannot* be cell autonomous. More precisely, if two cells share a common junction, it becomes difficult to conciliate how a cell undergoing EMT can do so without affecting its partner. In other words, if EMT occurs sparsely as is the case in the onset of carcinogenesis, it is not possible to predict whether the epithelial phenotype of the neighboring cells will be maintained, nor if these cells will influence the induced cell to remain epithelial.

3 Autonomy, causality and geometry

3.1.1 Geometry and causality

From the preceding discussion, the induction of EMT by a given cause is expected to lead to effects that could strongly depend on the geometry underlying the distribution of that cause. In other words, the relative position of induced and non-induced cells matters. Interestingly, seminal experiments done at the very beginning of viral molecular oncology raised this point. The existence of a normalizing signal has been experimentally demonstrated in pioneering viral oncology work by Rubin where a monolayer of primary chick fibroblast is shown to be differentially transformed by the Rous sarcoma virus (RSV) depending on the presence or absence of surrounding normal fibroblasts [83, 84]. This effect, termed "phenotypic suppression", was also reported for the polyoma virus [85, 86]. In fact, it is well accepted that many tumors may exist within an organ and still remain unnoticed by the host. These are referred to as “occult cancers” [2] and are thought to occur much more frequently than is commonly recognized. How these tumors are refrained from progressing into overt cancer is not fully understood. Inferring the underlying “normalizing” processes implies studying the tumor as part of a specific microenvironment and not only based on its internal determinants.

The ambiguity of cellular individuality in epithelia is not problematic as long as adjacent cells exhibit symmetric properties and contribute equally to junction stability. In this ideal situation, the system undergoes small fluctuations due to intrinsic transcriptomic noise, but remains stable as these fluctuations are counter balanced. This has been described as ordered heterogeneity: “order in the large over heterogeneity in the small” [8, 87]. In cervical precancerous lesions, the major etiology involves infection by the human papilloma virus (HPV) [88]. However, it is clinically known that there is no univocal relationship between the cause (infection) and its expected effect (dysplasia, transformation): some infected cells are not dysplasic, while some apparently dysplasic cells seem to not be infected [89]. In addition, these precancerous lesions can either regress or evolve to actual tumorigenesis [90]. This situation illustrates the notion that there is no categorical link and determinism between a cell’s internal state and its corresponding phenotype, and that a deductive strategy is incorrect when trying to predict the outcome of a heterogeneous junction [5] (Figure 1.1, c). The collective nature of the epithelium leads to an absence of a univocal cause/effect relationship, and to a non-deterministic interpretation of the cellular events underlying epithelial phenotype formation.

These non-reductionist examples clearly illustrate the importance of considering cells in a collective context rather than based on their sole individuality. It also suggest the existence of emergent properties which can only be explained when the contribution of intercellular interactions are considered. Whether a heterogeneous situation evolves in one direction or the other remains unpredictable. The stakes of understanding what governs the fate of heterogeneous situations are numerous, and their comprehension would drastically influence our understanding of carcinoma as well as possible treatment methods.

3.1.2 Cellular fitness

Rubin laid the groundwork for the vast problem of cellular heterogeneity [91]. More recent experiments have shown in various models that the emergence of heterogeneity leads to cellular competition and cellular fitness, giving rise to winner and loser cells [92, 93]. Examples of competitive interactions promoting tumor formation show that precancerous winner cells can induce apoptosis in loser wild type cells in an epithelium [10, 94], or colonize the stem niche by eliminating the loser wild type cells in the case of hematopoietic stem cells [95, 96]. The same type of competition takes place to suppress aberrant proliferation of transformed cells by wild type cells. In this case the former are considered the loser cells, while the latter are the winner. Such a situation has been observed in *Drosophila* where wild type cells induce cell death and extrusion of knockdown mutant cells of dCsk, *Drosophila* C-terminal Src kinase [97]. This has also been reported in MDCK cells transformed by activated oncogenic Ras which are extruded from the monolayer when surrounded by wild type cells [98]. It thus appears that this competition may either go down the routes of phenotypic suppression or phenotypic propagation, and that the outcome of this competition depends on the relative fitness of individual cells and on their environment As a consequence, the outcome of a heterogeneous situation is not the result of a random process, but rather governed by possibly precise competitive rules to be discovered.

3.1.3 Analogies with physical concepts of phase transition

In cancer, cells can be schematically modelled as two main physiological states: normal cells, and cancer cells. Through processes such as EMT and MET, cells can transition from one state to the other in a very dynamic manner. In fact, these type of transitions bear some analogies with phase transitions in the physical sense. Matter exists in various distinct states or phases, and when driven by certain external factors referred to as *control parameters*, a phase transition may occur at a critical value of these external parameters. A simple example is water which starts to boil (at standard atmospheric pressure) when the temperature reaches 100°C, and changes from liquid to

vapor. Here, the stability of the water phase relative to the vapor phase decreases as temperature increases (or control parameter is varied): the liquid state is stable and vapor state is metastable (less stable but not necessarily unstable) below 100°C, whereas above 100°C, the situation is opposite. This is reversible since decreasing the temperature will cause the reverse phase transition. The major physical changes in the system under study are described as *order parameters* (density, polarization ...).

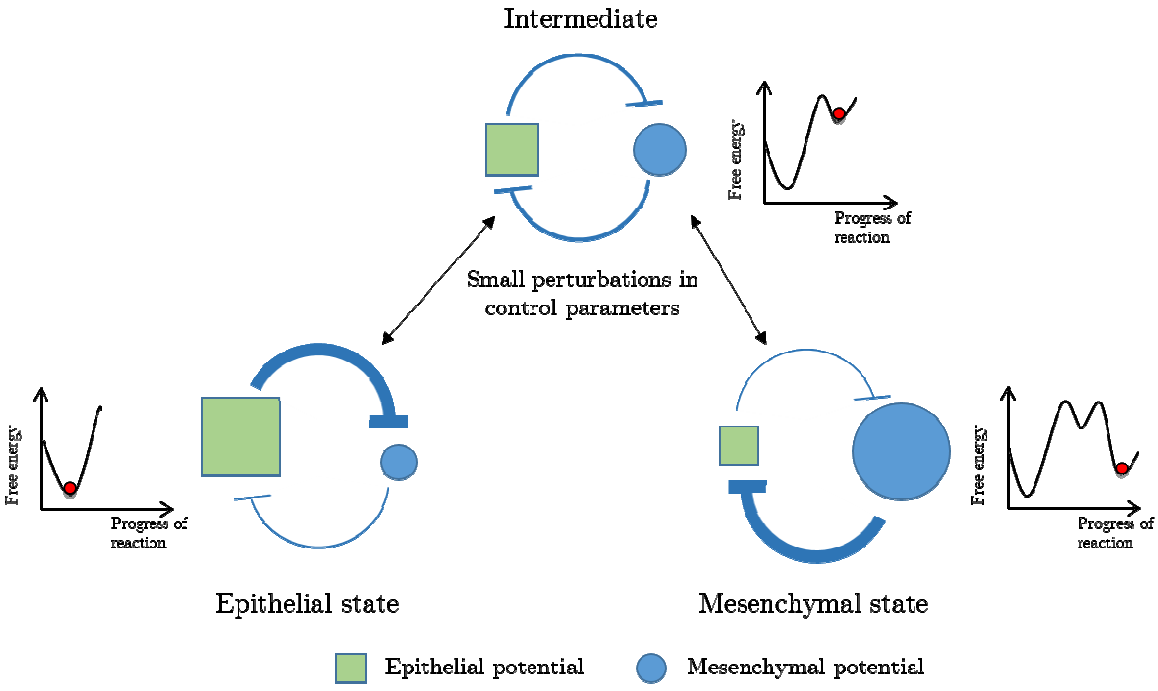


Figure 3.1 EMT as a dynamical phase transition

EMT and its subsequent reverse event, MET, allows for cells to exist either in an epithelial state or a mesenchymal state in a manner that is comparable to phase transitions. Going from the epithelial state to the mesenchymal state (or the reverse) requires to shift the relative stability of these two states by overcoming high free energy barriers due to the coexistence of the two phenotypes part of unstable hybrid intermediates. Due to this high instability, small changes in control parameters will cause the system to settle for one of the two stable states. Understanding what may be the control parameters regulating EMT is key when trying to predict the outcome of dysplasic and unstable intermediates that characterize cancer heterogeneity. Adapted from [10].

Applying the extensive body of knowledge regarding phase transition to cancer suggest that cancer onset will be extremely sensitive to external perturbations, and will occur at certain critical thresholds of sensitivity which can be measured [99] (critical density of mutated cells or cluster size). This prompted the setup of experiments where

key concepts of phase transitions could be used: nucleation barriers, metastability and critical points (Figure 3.1). In addition, it suggests that cancer onset is a reversible process similarly to a physical phase transition that may be reversed by a change in the value of its *control parameters*. This also has implications in treatment possibilities as it may highlight that local interventions which do not change the global state of the organism are not sufficient to reverse cancer progression. A global action instead could lead to cancer reversal [99].

4 Thesis rationale

The present Ph.D. research primarily aims at assessing the relative contributions of internal vs. external determinants underlying the stability, or instability of epithelia. The strategy is to create a situation where these two classes of determinants play antagonistic roles. This is achieved by inducing the destabilization of an epithelium, using a cause that can be delivered at the level of the single cells according to various geometries (Figure 4.1). We study if, and how, the geometry underlying the distribution of that cause modulates the cellular response. How the cause-effect relationship depends on the geometry should reveal important information about cell autonomy vs. interdependence within the epithelium. Practically, we use an *in vitro* cellular model of MDCK cells in which EMT induction *via* Snail1 activation has been placed under optical control.

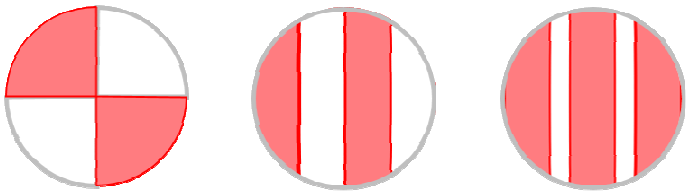


Figure 4.1 Heterogeneous induction of EMT

Local induction of EMT (red regions) in a monolayer of epithelial cells (white regions) grown in a culture dish (grey circle) will generate heterogeneous junctions. The fate of these junctions can be studied and a quantitative measure of cellular interdependence can be assessed as a measure of EMT nucleation or frustration in conditions where the geometric distribution of the cause is varied.

The Result section of this manuscript is organized as follow:

1/ Characterization of the cellular model for the induction of EMT, and validation of the chemical and photochemical activation of Snail1 with comparison to TGF-β.

2/ Establishment of essential readouts of EMT, and generation of a readout for the photoactivation event.

3/ Transcriptional analysis of spatially distributed heterogeneous EMT, and quantification of the response in a transcriptional state space.

4/ Generation of a novel algorithm for optical induction of EMT at the single-cell levels.

Chapter 2

II. Material and Methods

5 Cellular assays

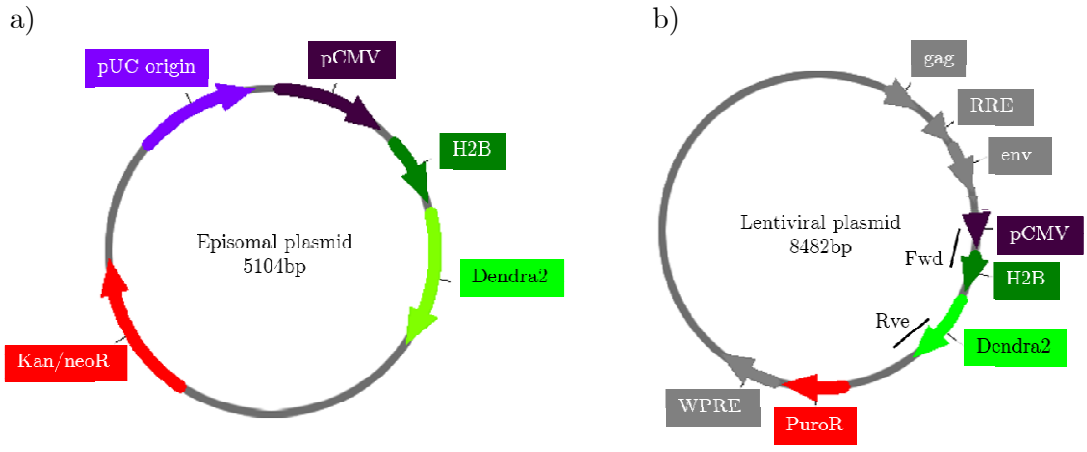
5.1.1 Cell Culture

MDCK cells were maintained in DMEM GlutaMAX (#61965-059, Gibco) supplemented with 10% FBS and 1% Penicilin/Streptomycin (Pen/Strep) (#15140122, Life Technologies). The MDCK-SnailERT2 cell line was a kind gift from Guy Lyons [100] (Sydney University, Australia). The cell line was continuously selected with 400µg/ml of Geneticin (G418) (#10131019, Life Technologies). The MDCK-SnailERT2-H2B-Dendra2 was further selected by adding 2µg/ml Puromycin (#P8833, Sigma Aldrich) under constant selective pressure. The medium was changed every 3 days and cells were passed at 80-90% confluency. Cells were maintained in culture in plastic cell culture flasks until passage ≈25 after which new cells were thawed. Cells used for experiments were platted in cell culture flasks, or on microscope glass coverslips (#S175223, Fisher Scientific, no.1), that were previously coated with poly-L-lysine (#P4832, Sigma). Briefly, a 1:1 volume solution of poly-L-lysine in PBS was prepared and spread on the surfaces to coat. The latter were place for 30min in the incubator and were then washed three times in PBS before cells were seeded. Coated flasks were prepared fresh and used immediately. Stock solution of poly-L-lysine was kept at 4°C and used within a few months after opening of the vial.

5.1.2 Generation of H2B Dendra2 stable cell lines

The pH2B-Dendra2 episomal plasmid was a kind gift from Xavier Darzacq, (Ecole Normale Supérieure, Paris) [101, 102]. The H2B-Dendra2 plasmid was then sub-cloned in the multiple cloning site of the pCDH1CMV-MCS-EF1-Puro lentiviral vector (#CD510B-1, System Biosciences) using In-Fusion® strategy (#639649, Clontech). Briefly, this method allows the fusion of DNA fragments into linearized vectors by recognizing a 15bp overlap between these two sequences. This 15bp overlap is generated

by designing primers that amplify, on the DNA fragment, the sequence of the vector corresponding to the site of insertion (Figure 5.1). The generated lentiviruses were packaged by co-transfection of this plasmid with pVSV (#12259, AddGene) and p8.74 (#22036, AddGene) using phosphate calcium in HEK293T cells. Medium was changed and collected every 24hrs for 3 days. The pooled supernatants were filtered through a 0,45 µm filter. Infectious titers were determined by infection of HEK293T cells using serial dilutions of the virus and FACS cell sorting of GFP positive cells. Subsequently, MDCK WT and MDCK SnailERT2 cells were infected with the H2B-Dendra2 lentiviruses at a multiplicity of infection (MOI) of 10. This value indicates the number of vector particles per cells used for transduction, and a MOI of 10 indicates that, ideally, each cell should be transduced by 10 viral particles. Using such a high MOI ensures that nearly all the cells of the culture are infected. Infected cells were then selected with 2µg/ml Puromycin. The cells generated (MDCK-WT-H2B-Dendra2 and MDCK-SnailERT2-H2B-Dendra2) were screened for their ability to respond to TGF-β and cyclofen (or not, in the case of the MDCK WT SnailERT2 H2B Dendra2 cells). Their morphology were comparable to that of the respective parental cell lines. Once established, the H2B-Dendra2 cell lines were subjected to continuous Puromycin selection.



c)

Forward primer:

5’ acctccatagaagattctagaatgccagagccagcgaag 3’

Reverse primer:

5’ gcggatccatttaaattcgaatttacttgtacacgccgc 3’

Figure 5.1 Episomal and lentiviral plasmids used for the establishment of H2B Dendra2 stable cell lines

The episomal plasmid (a) containing the H2B Dendra2 fusion protein under the control of a CMV promoter. The photoconvertible protein sequence was digested out and fused into a lentiviral plasmid (B) through In-Fusion® cloning strategy. The designed primers to generate the 15bp overlap are

shown (c). Letters in grey correspond to the flanking sequence of the linearized vector where the insert will be fused. Letters in dark green correspond to the sequence of H2B and letters in bright green to the sequence of Dendra2.

5.1.3 Immunofluorescence

Cells were grown on 18mm diameter glass coverslips (#48382-041, VWR) and washed twice on ice in cold PBS. Cells were fixed in room temperature (RT) 4% paraformaldehyde (PFA) for 15 min and washed 5 times with room temperature PBS (#15710, Euromedex). For staining internal proteins, cells were permeabilized using 0.25% Triton-PBS for 4 min at RT and coverslips were washed until the PBS drop no longer spread beyond the coverslip. Blocking was done for 30min using 0.25% (w/v) cold water fish gelatin (#G7765, Sigma) in PBS (blocking solution). Primary antibodies were incubated for an hour at RT in blocking solution, then washed 3 times for 10 minutes each with 0.125% gelatin in PBS (washing solution). Secondary antibody incubation was done using Alexa fluor antibodies and Hoescht for 45 minutes at RT in blocking solution. Finally cells were washed 3 times for 10 minutes each with washing solution. Coverslips were mounted on glass slides using Immunmount (Thermo Scientific), and stored away from light at -20C for several months. Cover slides were observed at the PICT-IBiSA imaging facility using an inverted scanning laser microscope (Zeiss) and 405nm, 488nm and 561nm lasers. Imaging was done with a 40x/1,3 Plan APO (#420762-9800, Zeiss), or a 63x/1,4 Plan APO (420782-9900, Zeiss) objective, both are oil immersion objectives. The system is equipped with a PMT GaAsp detection system. The various filters and dichroic mirrors used are summarized in the following table:

Filter Cube	Excitation range	Excitation filter	Dichroic mirror	Emission Filter
DAPI	UV	365	395	445/50
HE GFP	Blue	470/40	495	525/50
HE DsRed	Green	545/25	570	605/70

5.1.4 Antibodies

Primary antibodies: ZO-1: Invitrogen cat#339100, mouse monoclonal (at 1:100) or Invitrogen cat#61-7300, rabbit polyclonal (at 1:100). **Vimentin**: Dako cat#M0725 clone V9, mouse monoclonal (at 1:300). **E-cadherin**: BD laboratories cat#610181, mouse (at 1:100) or Invitrogen cat#13-1900, mouse monoclonal (at 1:100-1:200). **F-actin**: Phalloidin-TRITC.

Secondary antibodies: All from Thermo Fisher; Alexa Fluo® 488, Alexa Fluor® 555, Alexa Fluor® 633, Alexa Fluor® 647 (at 1:1000), **Hoescht:** 33342 ThermoFischer cat#H3570 (at 1:1000).

5.1.5 TGF-β induction

Cells were washed twice in 1x PBS and starved for 24hrs in starving medium: DMEM GlutaMax and 1% Pen/Strep (no FBS). Cells were then washed and incubated for 48hrs in normal medium containing 10ng/ml TGF-β (#P01137, R&D Sytems). Negative controls for the effect of TGF-β were done using the same volume of a 0.1% bovine serum albumin (BSA) solution. TGF-β stock solution was resuspended according to the manufacturer's instructions at a concentration of 2mg/ml, and stored as aliquots at -20°C.

5.1.6 Cyclofen induction

Cells were washed twice in RT 1x PBS, and placed in the incubator for 20min in starving medium containing 5µM cyclophen. Cells were then washed once in RT 1x PBS and incubated with normal full medium. Cells were observed 24hrs after cyclofen induction. Proper homogenization of cyclofen in the culture medium is critical for induction and can be best achieved in warmed medium due to limited solubility. The medium is heated to 37°C by placing an aliquot in the incubator. Cyclofen was then added to this warm medium and homogenization was ensured by up and down pipetting. This mixture was applied onto the cells. Negative controls for the effect of cyclofen were done using DMSO at the same volume. (Molecular weight (MW): 351,48 $g.mol^{-1}$). Aliquoted stock solution was kept for several months at 10mM concentration at -20°C.

5.1.7 Caged Cyclofen induction

Cells were washed three times in RT 1x PBS, and placed in the incubator for 1hr in starving medium containing 10µM caged cyclofen. Cells were then washed once at RT and full growth medium was added. Experiments were protected from light using aluminum foil to avoid undesired uncaging of the drug. The same issues of solubility exist with caged cyclofen, and the drug was handled similarly to cyclofen. UV illumination was done by shining light from the bottom of the dish. After illumination, medium was changed, and cells were incubated 24hrs. Dilution of the caged cyclophen in the starving medium was separately done in individual tubes to ensure proper homogenization of the drug prior to adding it to the cells. Controls for the effect of cyclophen were done using RT DMSO at the same volume. (Molecular weight (MW): 546,65 $g.mol^{-1}$). Aliquoted stock solution was kept for several months at 10mM concentration at -20°C. Powder

stock was kept at -20°C for years. Tubes were protected from light with aluminum foil to prevent undesired uncaging of the molecule.

5.1.8 Photoactivation

After incubation in caged cyclophen and washing in PBS, cells were subjected to UV illumination. For large patterns, or whole dish photoactivation, cells were illuminated using a 365nm and 700mA/190mW (Min) UV lamp (#M365L2-CA, ThorLabs). The intensity was measured using a powermeter, and the corresponding exposure time required to uncaged the cyclofen was calculated according to [103]. Cells were washed once in PBS after illumination and full medium was replaced. Cells were incubated for 24hrs. Controls for the effect of UV were also done. For small patterns designed at the single cell level, illumination was performed with a 405nm laser mounted on a spinning disk microscope. This is described in more details in Section 5.3.2.

5.1.9 Conditioned medium experiments

To test whether effectors of EMT could be present in the medium, conditioned medium experiments were carried out. The conditioned media were prepared from cells in culture in 6-well plates. Cells were either induced with DMSO (control) or with caged cyclofen in a homogeneous manner as described in Section5.1.6 and 5.1.7. After 24hrs of incubation the supernatants were collected giving rise to "epithelial" primed and "mesenchymal" primed media respectively. The latter were briefly centrifuged to remove cell debris. At this point, new cells were induced and photoactivated both homogeneously and heterogeneously. After UV illumination, cells were washed once in PBS but instead of replacing full medium, cells were incubated in the conditioned medium previously prepared. The "epithelial" primed medium was placed onto EMT induced cells whereas "mesenchymal" primed medium was placed onto non EMT induced cells. Cells were then incubated for 24hrs and then FACS sorted as described in Section 5.2.1.

5.1.10 Culture of 3D cysts

MDCK cells were grown in 3D in 8 chambers LAB-TEK (#055071, Dutscher) slides in medium containing 2% Matrigel (#354230, lot #3101587, lot#3101587, BD BioScience). Wells were humidified prior to seeding by washing with a small volume of PBS. Matrigel was thawed on ice to remain liquid. Each well was coated with 5µl of 100% Matrigel. The slide was then placed for 5min in the incubator at 37°C in order for the Matrigel to become a gel. Cells were prepared as a unicellular suspension of 30 000 cells/ml. In parallel, a 4% Matrigel solution in normal culture medium was prepared. A 1:1 mixture of these two solutions was done to obtain a final mix of 15 000 cells/ml in 2% Matrigel. 300µl of this mix was plated in each well of the 8 chamber slide. Medium

was changed every 2-3 days. Cysts were observed every two days with phase imaging and further grown for about a week until they reached a clear 3D conformation with a hollow lumen.

5.2 Biochemical assays

5.2.1 FACS and analysis

24hrs after UV illumination, cells were FACS sorted based on their relative GFP-like and RFP-like emission levels. Cells were washed once in PBS, trypsinized, collected and centrifuged. A recurrent observation is that UV illumination seems to make the trypsinization step less efficient, and longer trypsinization time was needed. This may be due to some cross linking effect of the UV light on the poly-L-lysine. The pellet was resuspended in 200-300µl of FACS buffer (PBS containing 2% FBS) and placed on ice. Tubes were protected from light avoid bleaching of the fluorophores. This suspension was then filtered through a 35µm cell strainer to remove cell aggregates. The FACS sorting step was done at the Flow Cytometry facility at the Curie Institute using a BD FACSARIAIII cell sorter. To establish proper gating of the fluorophore, single color samples were run, as well as unstained cells (MDCK GFP/RFP and MDCK WT respectively). Compensation was applied to get rid of any fluorophore overlap between channels. Sorting gates were placed such that the most GFP-like positive and most RFP-like positive cells only were collected. The region in between was discarded to avoid false positives and false negatives (Figure 7.10). Cells were collected directly in the lysis buffer provided by the RNA extraction kit (#79216, Qiagen). This ensured that the RNA from the cells were protected from degradation directly after the sort. Samples were extracted directly after the sorting or stored at -20°C if the purification was done on a separate day. In such case, samples were heated at 37°C for 15-30min prior to extraction in order to re-dissolve the salt contained in the lysis buffer and to obtain better purification yields.

5.2.2 RNA extraction and purification

For adherent cells in culture, culture dish was placed on ice and cells were washed twice in sterile RNAse-free 1x PBS. RNA extraction was done using the RNAeasy Plus Mini kit (#74134, Qiagen) following the manufacturer's instructions. Total RNA and purity was measured using a Nanodrop ND1000 Spectrophotometer. Extracts were stored at -80°C. Samples that were FACS sorted were much smaller in size than samples originating from cell culture dishes. As a result, RNA extraction was done using the RNAeasy Plus Micro kit (#74004, Qiagen), which, compared to the previous kit, allows

for a more efficient purification of small samples. Similarly, extracted RNA was quantified and stored at -80°C.

5.2.3 Reverse Transcription and cDNA synthesis

Reverse transcription (RT) reaction was performed to generate a cDNA libraries from the freshly prepared RNA extracts. The reaction was performed using the Biorad iScript kit (#170-8890, Biorad) starting from 500ng-1µg of RNA extract in a 20µl total reaction volume and following the manufacturer's instructions. The reaction mix was placed in a PCR machine with the following cycles: 5min at 25°C/30min at 42°C/ 5min at 85°C/ ∞ at 4°C. Samples were stored at -20°C.

5.2.4 qPCR

Quantitative PCR (qPCR) experiments were performed using freshly prepared cDNA. Because the RT reactions were done using the same quantity of starting RNA, cDNA quantities are expected to be more or less similar from one condition to the other. 1µl of cDNA was used in each well of the 96 well plate. Reaction mixture was composed of cDNA, primer pairs, SYBR reaction mix (#4367659, Life Technologies) and ddH_2O. qPCR 96 well plate was read by DNA Engine® (Biorad) and the following cycle was applied: 95°C for 10 min, (95°C for 15 sec, 60°C for 1min)x44 times and read plate. Melting curves from 55°C to 95°C (read every 1°C and hold 1sec) were generated. Conditions were performed in triplicates, and a negative control containing ddH20 instead of cDNA allowed to control for possible contamination. Primer amplification efficiency was measured by calculating the slope of a linear standard curve generated with varying known concentrations of RNA. Specific oligos were synthesized to study different EMT genes. The genes GAPDH, B2M and HPRT were used as reference/housekeeping genes and served as internal controls.

Fibronectin forward 5' GCAACTCTGTGGACCAAGG
reverse 5' CACTGGCACGAGAGCTTAAA [59]
Vimentin forward 5' TCTACGAGGAGGAGATGCGG
reverse 5' GGTCAAGACGTGCCAGAGAC [104]
Slug forward 5' AGCAGTTGCACTGTGATGCC
reverse 5' ACACAGCAGCCAGATTCCTC [105]
ZEB1 forward 5' TTCAAACCCATAGTGGTTGCT
reverse 5' TGGGAGATACCAAACCAACTG [59]
E Cadherin forward 5' AAGCGGCCTCTACAACTTCA
reverse 5' AACTGGGAAATGTGAGCACC [59]
GAPDH forward 5' CATCACTGCCACCCAGAAG
reverse 5' CAGTGAGCTTCCCGTTCAG[59]
HPRT forward 5' TTATAGTCAAGGGCATATCC
reverse 5' AGCTTGCTGGTGAAAAGGAC [106]

B2M forward 5' TCCTCATCCTCCTCGCT
reverse 5' TTCTCTGCTGGGTGTCG [106]

5.2.5 Analysis of qPCR data

Real-time PCR, also known as qPCR, allows to quantitatively determine the amount of a target sequence or gene that is present in a sample. Data generated by qPCR consists in amplification curves that were analyzed using the efficiency calibrated method [107]. The curves represent the amount of the target sequence detected after every cycle of amplification. The threshold cycle (Ct) is the cycle number for which the amplification curve meets a threshold line that is fixed at about 63% of the maximal amount of DNA has been reached. Amplification efficiency (E) was calculated for each oligo pairs by measuring the doubling rate of a known amount of starting DNA, and ΔCt was calculated for each gene (target vs. reference) as $\Delta Ct = Ct_{control}\text{-}Ct_{treatment}$. The ratio of target gene expression in treatment versus control is then derived from the ratio between target gene efficiency (E_{target}) to the power of target ΔCt (ΔCt_{target}) and reference gene efficiency ($E_{reference}$) to the power of reference ΔCt ($\Delta Ct_{reference}$) [107]. This is summarized in the equation:

$$Ratio = \frac{(E_{target})^{\Delta Ct_{target}}}{(E_{reference})^{\Delta Ct_{reference}}} \quad \textbf{(1)}$$

Where $\Delta Ct_{target} = \Delta Ct_{control} - \Delta Ct_{treatment}$
and, $\Delta Ct_{reference} = \Delta Ct_{control} - \Delta Ct_{treatment}$

$E^{\Delta Ct}$ was calculated for all genes as well as for the three reference genes using the calibrated value of E. The $E^{\Delta Ct}$ for GAPDH, B2M and HPRT were averaged together and the resulting $E^{\Delta Ct}$ was used as the denominator in Equation ((1).

5.3 Tools for optical illumination

5.3.1 Chatterton tape masks

For large scale heterogeneous photoactivation, Chatterton (electrician's tape) mask were designed and placed on the bottom of the cell culture well. To ensure precision, the masks were cut out using a numerical cutter (#CE5000-40, Graphtec Craft Robo Pro) which allows a resolution of about 200μm. The tape was placed into as spectrometer to verify that it was impermeable to UV light and could thus serve as a reliable mask. Patterns were designed on Illustrator.

5.3.2 Development of *Selective Illumination* (Coll. Thomas Walter, Curie Institute)

The selection of cells according to specific neighbourhood relationship rules (with free parameters to be set prior to the experiment), is done in a fully automatic way by an open-source software that was specifically designed for this study. The objective was to probe for specific neighbourhood conditions, i.e. inducing the phenotype in clusters of neighbouring cells separated by cells in which the phenotype was not induced. The free parameters are therefore the cluster size, i.e. the number of neighbour cells in which the phenotype is induced and the distance between clusters, i.e. the number of non-induced cells separating induced clusters. The method to automatically select cells according to these rules consists in 3 steps: (1) segmentation of the cell nuclei, (2) building a neighbourhood graph and (3) selection of cells according to the chosen neighbourhood parameters. The three steps are modular; it is therefore possible to replace, for instance, the specific segmentation method adapted to the particular imaging method used in this project by another segmentation procedure without any consequence on the other steps of the method. The software was implemented in Python and made use of the image analysis toolbox scikit-image [108]. A graphical interface was also implemented in order to allow to easily change parameters of the neighbourhood conditions to be tested (Figure 5.2). The software (including the GUI) is publicly available on github (https://github.com/ThomasWalter/SelectiveIllumination.git).

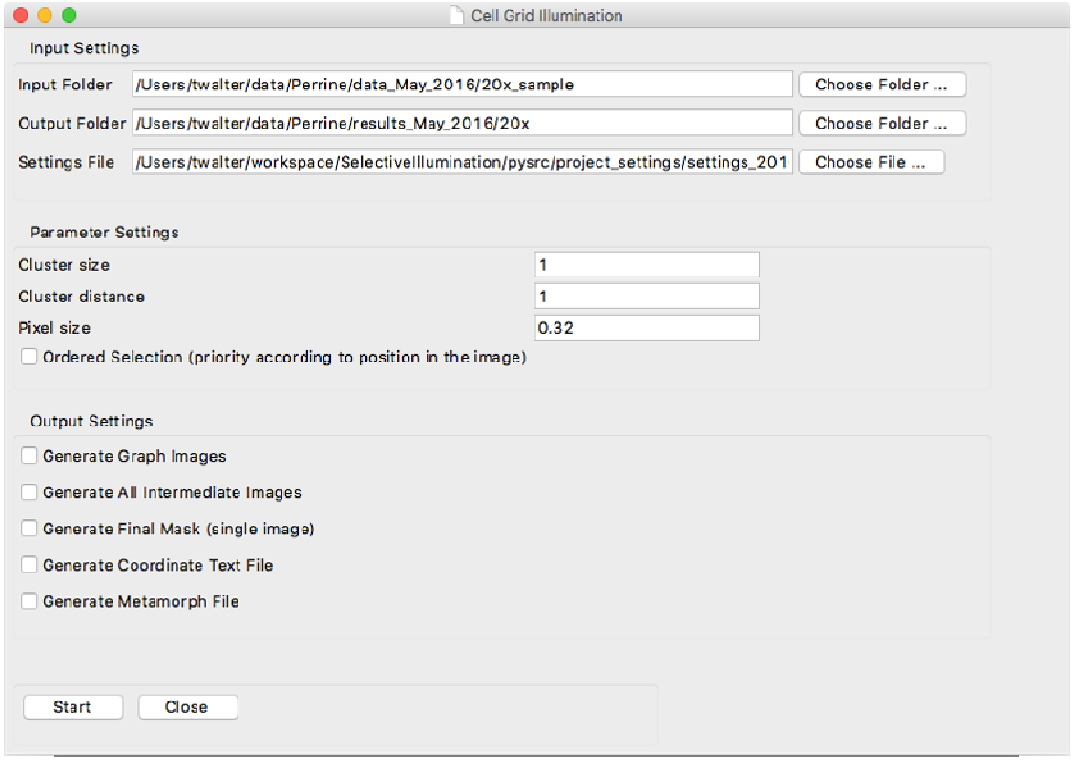


Figure 5.2 Graphical User Interface of the software

The output of *Selective Illumination* is a list of positions in a .stg format that is sent back to the microscope, ordering the FRAP head to turn on the laser only on these specific locations. The position defining a given cell corresponds to the center of its nucleus. These experiments were conducted on an inverted spinning disk microscope running under MetaMorph, using a 405nm laser beam.

5.3.3 Statistical analysis

All the data was analyzed using OriginPro 9.1. P-value for determining statistical significance was calculated by a one-tailed one-sample Student's t-test, and p values less than 0.05 were considered significant. Different levels of significance are shown with asterisk on the graphs: *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$.

Chapter 3

III. Results

6 Biological Model

At the time of start of this thesis work, the project was at its most early stage of brainstorming. As a result, planning, designing and generating the necessary tools constituted a significant amount of work. Given the previous introduction, the biological model needed to address the posed question had to fulfill a certain number of requirements. Indeed, to study intercellular communications in the scope of understanding some aspects of the onset of carcinogenesis implied that the chosen model should be comprised of cells that are physically interacting. Furthermore, it is critical that the phenotype of the cells remain stable over time and that they do not undergo spontaneous differentiation. In this biological model, it then becomes feasible to use optogenetics to locally implement a perturbation causing a change in phenotype that can be studied over time. This chapter covers the different tools that have been setup as well as the most significant results obtained.

6.1 Stable cell line

6.1.1 Choice of the cell line

The main requirement for the cellular model in this project was that cells be epithelial and that this phenotype be stable over time. The canine epithelial cells, Madin-Darby Canine Kidney cells (MDCK), came as a very natural choice partly because they constitute a very well characterized epithelial model for the study of EMT, epithelial polarity [109] and plasticity [50, 105]. This same cell line had been previously used in the laboratory for the study of cadherin endocytosis and the study of junction dynamics [110]. In addition, the team of Guy Lyons had generated a stable cell line of MDCK cells

expressing a key transcription factor of EMT, thus making available a robust tool for the project [100].

Using cells of human origin is of course very relevant for the pathology-related interpretation of any possible results. However, after thorough examination of the literature, no ideal human epithelial cell line was found. For instance, human pancreatic epithelial cells (PANC-1) as well as human mammary epithelial cells (HMLE) were considered, but EMT induction in both cell lines proved to be non-trivial to elicit a detectable effect. In addition, there is evidence that the epithelial phenotype in these cells is not stable over time and can in fact, in certain cases, undergo spontaneous EMT as a result of up-regulation of autocrine TGF-β signaling[111].

6.1.2 Choice of the model

The cell line used, referred to as MDCK-SnailERT2, was stably transfected with the fusion protein SnailERT2. As mentioned in the introduction, Snail1 is a key inducer of EMT. Activation of Snail1 has been shown to be sufficient to trigger effective EMT [112-114] and will be used in this project as the mediator of EMT in individual cells. The transcription factor is part of a fusion protein with ERT2 which is linked to the N-terminal of the full sequence of Snail1[100]. ERT2 is a modified version of the estrogen receptor which retains only the hormone binding domain of the estrogen receptor (ER) [103]. This fusion protein is thus ligand dependent, such that in the absence of ER ligand, it is sequestered in the cell cytoplasm by chaperones such as hsp90. Addition of the ligand triggers the unbinding of the fusion protein from these chaperones and allows the fusion protein to enter the nucleus in its active conformation. Modifications on the ERT2 moiety confers more selectivity to the artificial steroid tamoxifen (4-hydroxy tamoxifen, or 4-OHT) compared to its natural counterparts [115, 116].

Because ERT2 is part of a fusion protein, its nuclear translocation also brings along Snail1. As mentioned previously, there is a strong correlation between nuclear localization of Snail1 and its activation. In other words, nuclear translocation of ERT2 as a result of ligand binding, causes Snail1 translocation and thus activation as well as induction of EMT through the repression of E-cadherin promoter [117]. With this system, it is possible to trigger EMT in MDCK-SnailERT2 by the addition of tamoxifen.

Given this, it is however important to state that the true compartmentalization of the SnailERT2 fusion protein is most likely not straightforward. In fact, it has been shown that certain hormone binding domains, part of ERT2, can be localized to the nucleus in the absence of the ligand and remain transcriptionally inactive because of ongoing binding of hsp90 and other mechanisms [118]. In addition, it is difficult to anticipate how the import/export regulation of Snail1 described earlier holds true in the

context of a fusion protein that is larger in size and subjected to possible antagonistic ERT2 regulation mechanisms. In fact, preliminary experiments revealed that the SnailERT2 fusion protein was observed to be in the nucleus in the absence of any ligand (data not shown). In other words, the MDCK-SnailERT2 cell line has been shown to respond to its specific ligand [100], but the exact regulation and localization of the fusion protein is not fully understood: protein localization cannot be used as a readout of EMT induction.

The team of Ludovic Jullien at ENS Paris has studied the interaction of 4-OHT with the modified estrogen receptor and developed an analog called cyclofen. The rationale and interest for this compound will be explained in latter section of this manuscript [103].

6.1.3 Validation of the cell line for the induction of EMT

Experiments were first done to validate the suitability of the chosen cell line for the induction of EMT, as well as to verify the functionality of both the fusion protein and the ligand as a potent inducer of EMT. Cells were either stimulated with transforming growth factor (TGF-β) for 48hrs, the gold standard for the induction of EMT, or with cyclofen for 24hrs. In both cases cells showed a change in morphology compared to the control condition (Figure 6.1, A-C). In the TGF- β condition, cell elongation was very obvious and cell appeared less densely packed. Cyclofen stimulation causes a loss in the lining of the epithelium's smoothness, or actin cable [119], and apparition of spiky edges, or star-like morphology, which is characteristic of the expected effect. The morphological change is however less pronounced than with the TGF- β. This can be explained by the fact that cyclofen acts only on Snail1 which is a downstream effector of the TGF- β pathway and as a result, its action is much less broad than TGF-β.

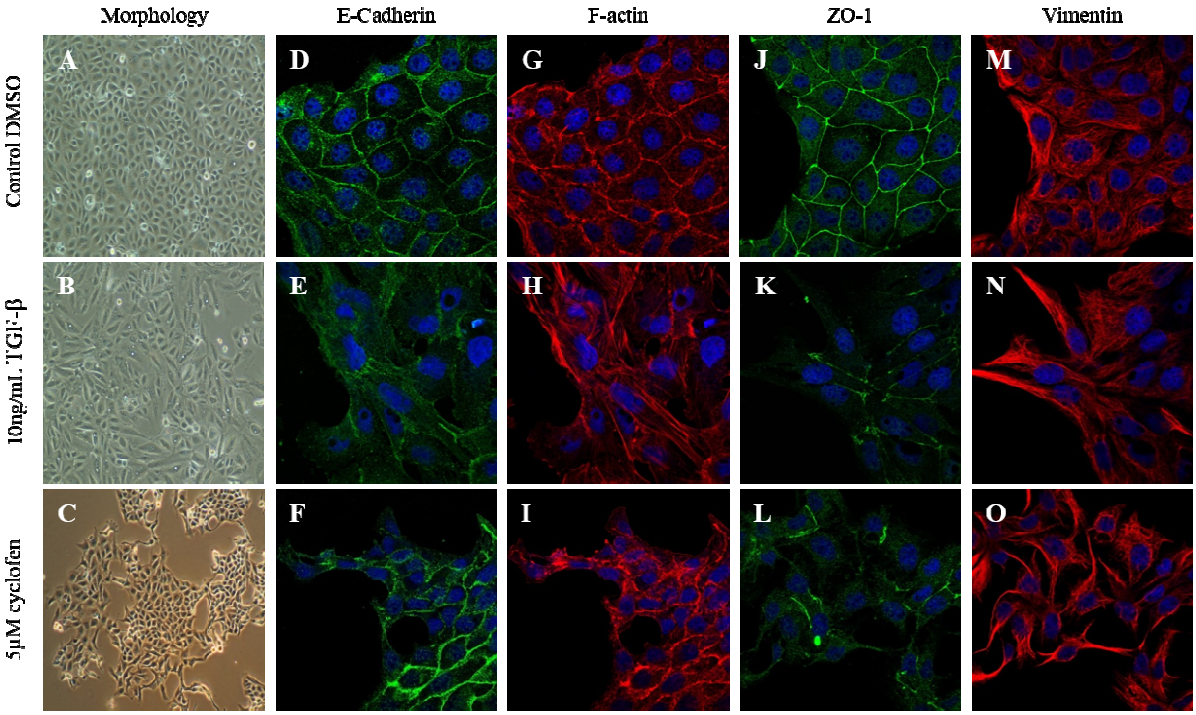


Figure 6.1 EMT induction in MDCK-SnailERT2 cells

Morphological images (A-C) and immunofluorescence stains for the indicated markers (D-O) in MDCK Snail-ERT2. Cells were imaged after 48hrs of DMSO (top row) or 10ng/ml TGF-β (middle row); cells were imaged after 24hrs following 20min incubation of 5μM cyclofen (bottom row). Cells were stained for the adherent junction marker, E-Cadherin (D-F) as well as for the tight junction marker ZO-1 (J-L). Cytoskeletal organization was assessed by both F-actin (G-I) and Vimentin (M-O) markings. Nuclei are counter-stained with DAPI (blue). (Detection channels: blue 405nm, green 488nm and red 635nm)

6.1.4 Choice and optimization of EMT markers

To further assess the response to TGF-β and cyclofen, immunofluorescence experiments (IF) were performed. Given that the cell line used is of canine origin, the choice of potential EMT markers was limited by the available functional antibodies in this species. Optimization of the IF conditions was done to enhance the signal generated by these antibodies. A large number of markers were stained on these cells, but only a few were useful. As a result, the morphological observations of TGF-β and cyclofen induction were paralleled with staining of four different markers of EMT. E-cadherin, as mentioned previously, is considered a fundamental marker of EMT and reveals the presence of adherens junctions [110]. Zona Occludin 1 (ZO-1) marks tight junctions [120]. These two markers are considered epithelial markers, and their expression is expected to decrease and/or relocalize during EMT. On the other hand, vimentin which is the major cytoskeletal component of mesenchymal cells and F-actin which polymerizes into stress

fibers when the cell elongates and migrates, are used here as mesenchymal markers. Other well established markers of EMT were also tested. More specifically, fibronectin, α-smooth muscle actin (α-SMA) and tubulin were considered, but the available antibodies were not adequate to detect a specific signal in MDCK cells. The IF experiments show that E-cadherin seems to be disengaged from the junctions but remains juxta-junctional as can been seen by the thickening of the signal. This is also accompanied by some punctuate cytosolic redistribution. ZO-1 on the other hand is completely lost from the cell junctions during TGF-β induction and cyclofen stimulation (Figure 6.1 D-F; J-L). In addition, apparition of actin stress fibers and enhanced vimentin signal suggests that cells have elongated and gained motility after cell treatment (Figure 6.1 G-I; M-O). The nuclei of the cells were counterstained with DAPI to appreciate the increased distance between the nucleus before and after treatment, suggesting that the epithelium is less densely packed.

As a side note, ZO-1 seems to be the clearest marker of EMT because of a strong modification of its staining upon EMT induction but, it raises the following issue. The localization at the junctions makes it impossible to tell whether both cells contribute equally or not to the staining pattern. To individualize the detection of the EMT response, a non-junctional marker is preferable.

The same experiments were done on wild type (WT) MDCK cells. Again, TGF-β induced strong cellular elongation seen both in morphology and in IF. On the contrary, cyclofen induction had no effect on the cells as they retained their epithelial phenotype, seen both morphologically and through the localized staining of both E-cadherin and ZO-1 (Figure 6.2). This control experiment proves that EMT can be induced in MDCK-SnailERT2 cells and that the response requires the SnailERT2 construct.

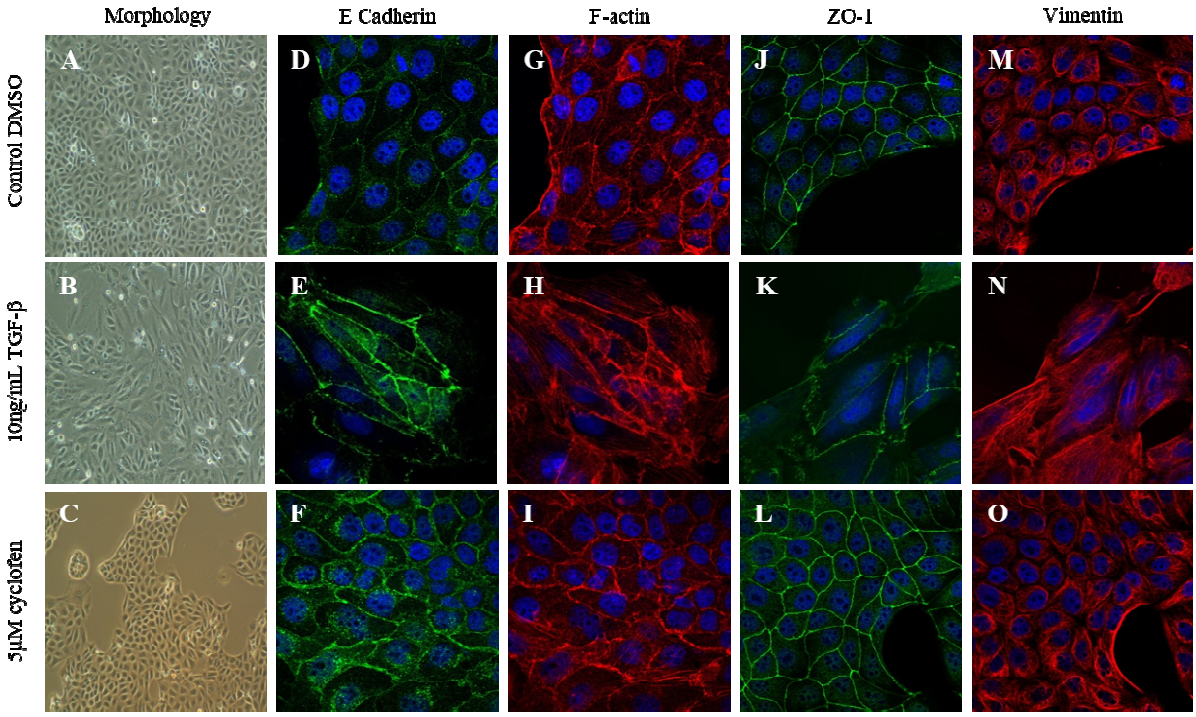


Figure 6.2 EMT induction in MDCK WT cells

Morphological images (A-C) and immunofluorescence stains for the indicated markers (D-O) in MDCK WT cells. Cells were imaged after 48hrs of DMSO (top row) or 10ng/ml TGF-β (middle row); cells were imaged after 24hrs following 20min incubation of 5μM cyclofen (bottom row). Cells were stained for the adherent junction marker, E-Cadherin (D-F) as well as for the tight junction marker ZO-1 (J-L). Cytoskeletal organization was assessed by both F-actin (G-I) and Vimentin (M-O) markings. Nuclei are counter-stained with DAPI (blue). (Detection channels: blue 405nm, green 488nm and red 635nm)

6.2 Quantification of the biological response

6.2.1 qPCR analysis of TGF-β and cyclofen induced cells

Responses to the different inducers of EMT has been assessed by morphological images as well as by IF. However, these analyses are mostly qualitative and the response requires a quantitative method in order to further appreciate the noticed changes. For this, quantitative PCR (qPCR) analysis was performed on the response of the SnailERT2 cellular model to the well-established EMT inducer TGF- β. This analysis will serve as a positive control for the expected response of the cyclofen-induced EMT. It is important to keep in mind that the signaling pathway of TGF- β is much broader and more upstream than that of cyclofen, so this positive control is the best possible reference available for what is expected with cyclofen. Five different EMT markers were assessed

and quantified using the efficiency calibrated method [107]. Ct cycles of each gene were normalized to the Ct cycles of GAPDH, B2M and HPRT, three housekeeping genes. As anticipated, results show that both MDCK-SnailERT2 and MDCK WT cells have similar responses to TGF-β (Figure 6.3). This is expected since TGF- β acts upstream of Snail1, rendering these two cell lines comparable in this case. This result also shows that the addition of the fusion protein in the MDCK-SnailERT2 does not affect the cells and their endogenous TGF-β response when it is kept inactive in the absence of its specific ligand. More precisely, Fibronectin, Slug, and ZEB-1 all increase significantly in response to TGF- β. Although E-cadherin and Vimentin levels do not vary in a significant manner, it can be precautiously noted that their relative fold change occurs in the anticipated direction.

Transcriptional levels in cells treated with tamoxifen and with cyclofen was then assessed. Because cyclofen is a novel molecule for the induction of EMT, the tamoxifen response will serve as more relevant positive control than the response obtained with TGF- β. Analyses show that both responses are similar in MDCK-SnailERT2 which confirms that cyclofen is an analog to tamoxifen in its ability to bind to and activate the ERT2 moiety (Figure 6.4, a). In addition, E-cadherin, Slug and ZEB-1 levels change significantly whereas Vimentin and Fibronectin do not. Similarly to Figure 6.2, experiments were carried out on MDCK WT cells to verify that the measured effect is in fact mediated by the fusion protein. In this case, none of the studied markers are significantly affected by tamoxifen nor cyclofen. This suggests that the likely presence of endogenous estrogen receptors (ER) in MDCK WT cells are not sensitive to tamoxifen. Furthermore, it confirms that the MDCK-SnailERT2 cells are responding to the addition of the ligand through specific ERT2 mediated activation.

Cells stimulated with TGF-β were observed and analyzed after 48hrs of stimulation. This time window is commonly used in the literature in various cell types [121-123] and proved to be sufficient in MDCK cells to induce a notable EMT event. The time window used for the induction of cells in cyclofen or caged cyclofen was set to 24hrs. This is based on published literature on the activation of ERT2 fusion proteins with the analog of cyclofen, tamoxifen [124, 125]. Here, induction of EMT shows that the parameters established for tamoxifen are also valid for cyclofen.

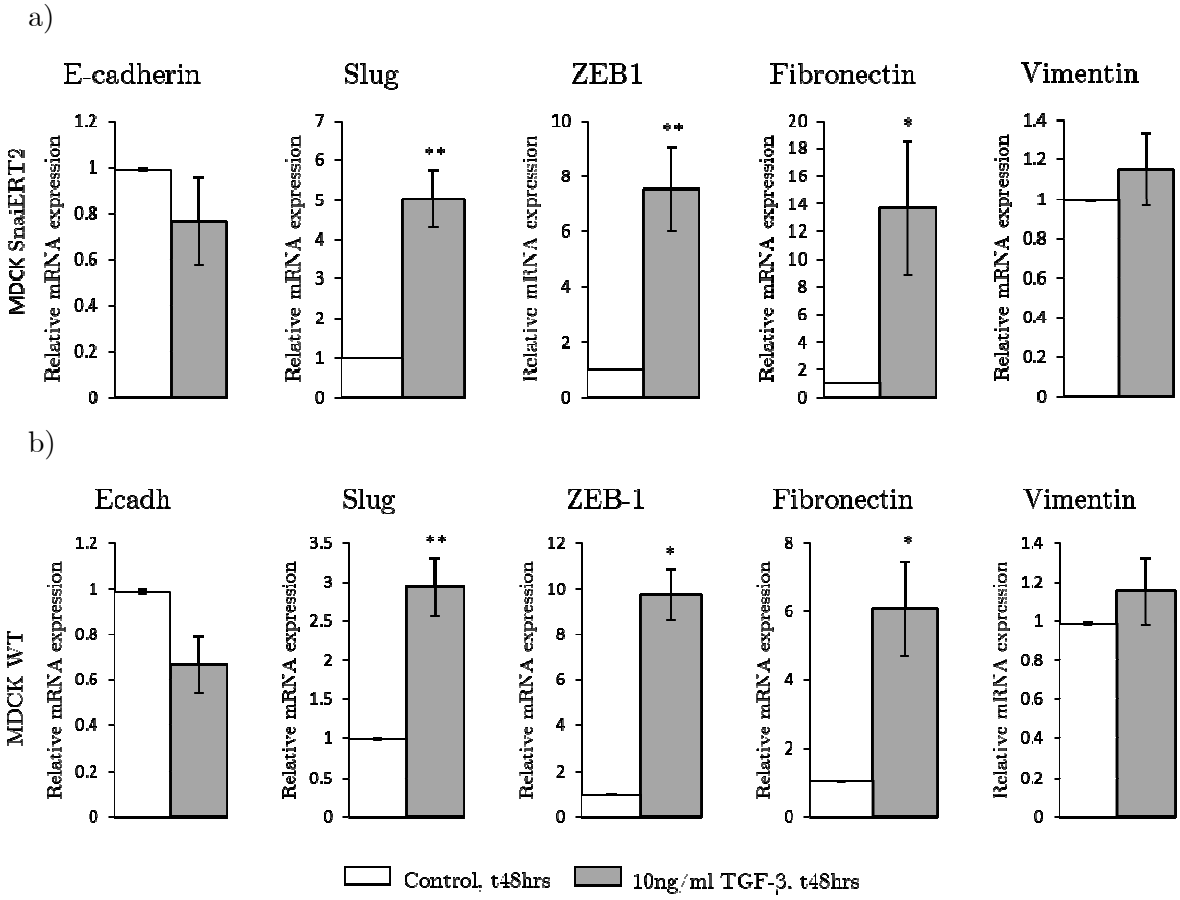


Figure 6.3 qPCR analysis of TGF-β induced EMT in MDCK-SnailERT2 and WT cells

Cells were stimulated for 48hrs with 10ng/ml TGF-β and lysed for RNA extraction. qPCR analysis reveal significant changes in expression of ZEB-1, Fibronectin, Slug. Levels of Vimentin and E-cadherin remained unaffected. Both MDCK-SnailERT2 (a) and MDCK WT (b) respond similarly. Error bars represent SEM, n=4, $p \leq 0.05$

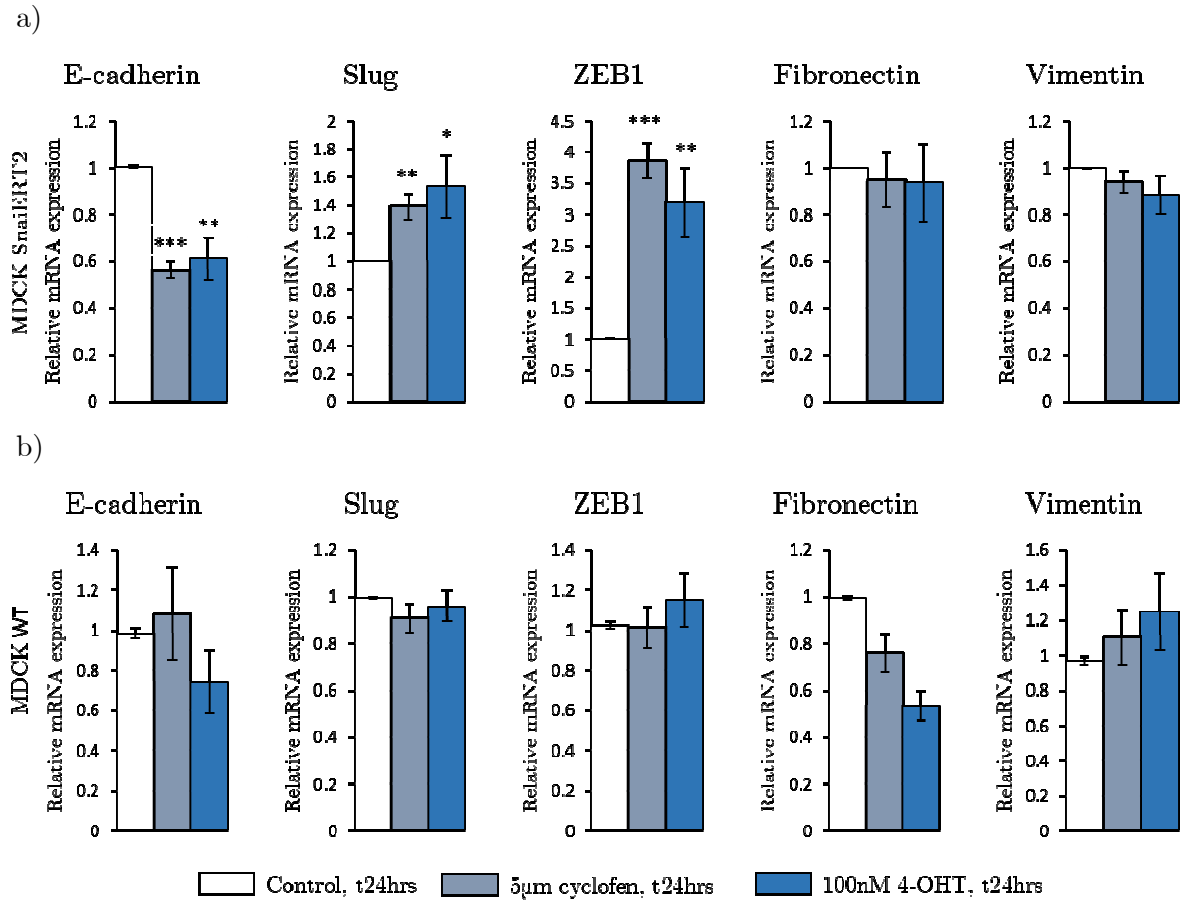


Figure 6.4 qPCR analysis of cyclofen and tamoxifen (4-OHT) induced EMT in MDCK-SnailERT2 and WT

Cells were stimulated with either 5µm cyclofen or 200nM 4-OHT for 24hrs and lysed for RNA extraction. qPCR analysis reveal similar changes for both drugs and significant changes in the expression of ZEB-1, Slug and E-cadherin in MDCK-SnailERT2 (a). Levels of Vimentin and Fibronectin remained unaffected. Experiments were also done on MDCK WT as a negative control (b), and result show no changes in expression of these markers in this cell line. Error bars represent SEM, n=4, $p \leq 0.05$.

7 Optogenetic approach

The goal of this thesis project is to study how the epithelium responds to heterogeneous activation, or local EMT induction in a manner similar to the onset of carcinogenesis. To address this question, the biological model requires a way of inducing local EMT. For this, the previously described cellular model was coupled to an optogenetic approach which will be described in the following part. Before defining the details of this method, it is necessary to introduce optogenetics.

7.1 What is optogenetics?

7.1.1 Early use of optogenetics

The term "optogenetics" was proposed by Karl Deisseroth, pioneer of this booming field [126] less than a decade ago. He defines this method as "the combination of genetic and optical methods to achieve gain or loss of function of well-defined events in specific cells of living tissues" [127]. Historically, optogenetics was developed to study neuronal circuits that require high spatial and temporal resolution as well as cellular specificity [128]. This was done through the introduction of microbial opsins in biological models via recombinant DNA methods. These naturally occurring light sensitive transmembrane proteins are found in various organisms, ranging from microbes to primates. When illuminated by light of the correct wavelength, opsins can be transiently activated or inhibited (depending on the type of opsins used) to induce ion channel opening/closing or changes in the membrane potential of a neuron [128, 129]. Building on the knowledge acquired with applications of excitable membranes, optogenetic capabilities were further extended to all possible cell type, including non-excitable tissues, through the use of genetically encoded light responsive proteins [130]. Interestingly, the term "optogenetics" in the first sense does not mean the control of genes by light but rather refers to the fact that light drives channel activity through indirect gene regulation.

7.1.2 Optogenetic systems in non-excitable tissues

In fact, proteins that change conformation in response to light have been engineered to regulate basic cellular functions such as protein localization or protein-protein interactions [131]. The most widely used photoreceptors are PHYTOCHROME B protein (PHYB), CRYPTOCHROME 2 protein (CRY2) and LOV domains. Very briefly, phytochrome domains are plant red/far-red light sensing photoreceptors that will undergo a reversible *cis-trans* isomerization when illuminated. This light dependent conformational change can be used to induce binding to or dissociation from downstream interacting proteins [131, 132]. Cryptochromes also come from plants but are UV-A/blue sensitive. The light-induced conformational change allows the binding of calcium and integrin-binding protein 1 (CIB1). CRY2-CIB1 heterodimerization is robust and rapid and is commonly used to recruit specific proteins to particular regions in the cell [130, 132]. LOV domains are somewhat similar to cryptochromes. Their small size allows different utilization of the conformational change: auto inhibition relief, recruitment of signaling domains or regulation of gene expression through DNA binding [130].

Overall, these photoreceptors, when attached to the protein of interest, can serve as powerful tools to induce rapid, reversible, specific and local cellular function alterations. Modification of gene expression, activation through clustering of signaling proteins, sequestration based inhibition, inhibiting/activating conformational changes and protein recruitment and localization are so many ways these powerful proteins can be used [133].

Finally, optogenetic approaches can also rely on the use of photoactivable or photoconvertible fluorescent proteins such as Dronpa [134]. When illuminated, these proteins not only undergo a conformational change that modify their quaternary structure and cause dimerization, but their fluorescence also changes. In its dimerized state, Dronpa presents a GFP-like fluorescence. This robust system can be used by fusing Dronpa to the amino and carboxyl group of a protein of interest. Activation of Dronpa by light will cause it to dimerize, thus possibly inhibiting protein function while activating the fluorescent state. Protein activity can then be assessed by the parallel increase in Dronpa fluorescence [134].

With this in mind, it was a very natural choice to opt for an optogenetic approach in this project, specifically for inducing heterogeneous patterns of EMT in a monolayer of epithelial cells. The system used here will be described in the following sections.

7.2 Inducible cellular model of EMT

7.2.1 A photoactivable ligand

Among the few systems available to chemically activate gene expression, the tamoxifen/ERT2 system is widely used. With the goal of using this receptor-ligand pair under optical control, the team of Ludovic Jullien set out to synthesize a photoactivable version of tamoxifen [103]. This truly optogenetic tool gives access to the control of virtually all the tamoxifen/ERT2-based methods.

In the previous section, cyclofen was shown to be a robust analog of tamoxifen (4-OHT) as seen by EMT induction both morphologically and quantitatively. The interest in using this analog rather than 4-OHT comes from the fact that its structure allows for the synthesis of a photoactivable version of cyclofen. This consisted in adding a caging group to the active tamoxifen, thus preventing it from being able to bind ERT2, its specific receptor. In theory, the protective group could be irreversibly broken off by exposing the molecule to ultra violet (UV) illumination (365-405nm). However, the structure of tamoxifen is susceptible to photoisomerization and photodegradation upon UV illumination under the required uncaging conditions. To overcome this, the analog

cyclofen was synthesized, and its structure allowed to generate a caged version, caged cyclofen, for which neither isomerization nor degradation were a problem. The caging group used is 4,5-dimethoxy-2-nitrobenzyl alcohol [103] (Figure 7.1). In addition, the hydrophobic and neutral structure of caged cyclofen allows it to passively diffuse through the cell membrane. When released, the active compound is believed to be retained into the cell as protonation of the amine functional group takes place at cellular pH, rendering the molecule more hydrophilic. This new compound was tested on the basis of its ERT2 binding capacities, as well as uncaging properties, and revealed that caged cyclofen, when exposed to UV illumination, mimics both cyclofen and tamoxifen actions [103]. It is thus a valid tool for the optogenetic induction of EMT in MDCK-SnailERT2 cells.

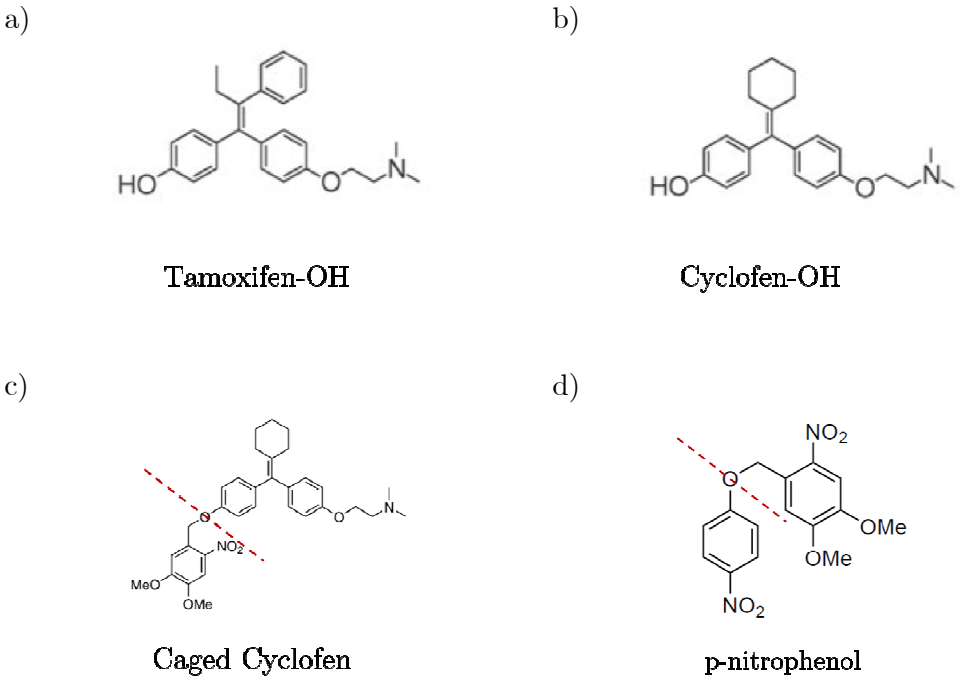


Figure 7.1 Chemical Structure of ERT2 ligands

Chemical structure of a) Tamoxifen-OH, b) Cylofen-OH and c) Caged cyclofen and d) p-nitrophenol. In c) and d) the caging group is a 4,5-dimethoxy-2-nitrobenzyl alcohol moiety. The red line highlights where the protective group is broken off under UV illumination. Adapted from [103].

With such a system, it is now possible through local delivery of light, to specifically trigger EMT in individual cells that have been illuminated. This irreversible optogenetic approach is the main tool used throughout this work to generate patterns of heterogeneous EMT within an epithelium.

Why use optogenetics? One could argue that optogenetics may not be required to study epithelial stability in the case of EMT and that co-cultures of EMT inducible and EMT non-inducible cells would be sufficient to recreate situations of heterogeneity. However, with this approach, there is no mean of dictating the spatial distribution of the heterogeneity thus limiting the control of various parameters (distance with the next induced cell, cellular density). In addition, preliminary experiments of the kind were done at the onset of the project, and it was noted that randomly seeded co-cultures of MDCK WT and un-induced MDCK-SnailERT2 cells do not mix properly, and sort themselves out over time to form separate homotypic domains or aggregates. In fact, intercellular junctions between homotypic cells differ in morphology from heterotypic junctions observed at the boundary of two slightly different populations. This has been observed by others and is explained by the fact that different populations of epithelial cells can recognize differences with their counterparts and fail to establish normal intercellular junctions [98, 135]. To overcome this, it is of utmost importance that all the cells in the system be initially "equivalent" in order to generate a stable homogeneous monolayer, where all the cells are contributing equally to junction dynamics. As such, the optogenetics approach allows, once this layer is formed, to locally deposit the perturbation in individual cells that are already physically engaged in truly homotypic junctions.

7.2.2 Theoretical calibration of photoactivation (L. Jullien)

In order to properly use the caged cyclofen, calibration of the kinetics of cyclofen release from its caged precursor was done. The main goal is to determine the required uncaging time to release the active cyclofen and observe the morphological effect. Two parameters that can be varied are illumination power and exposure time. Given that the wavelengths used (365-405nm) can generate some toxicity in the cells, it is important not to go beyond the requirements needed for the uncaging reaction to occur.

This theoretical calibration was done according to the work of Sinha. et al [103]. To relate the fluence of the illumination (J/s.cm^2) to its uncaging effect, one needs to know the following elements:

1) The energy of a single photon at 365nm

$$E_{375nm} = \frac{h\,x\,c}{\lambda} = 5.44 \times 10^{-19} J \quad \textbf{(2)}$$

Where h is Planck's constant and c is the speed of light

2) The absorption cross section which measures the absorption efficiency of caged cyclofen. Because the quantum yield of the uncaging process is not unity, the relevant parameter is thus the uncaging cross section [103]

$$\varepsilon_{uncaging} = 22000 cm^2 mol^{-1} = \ 3{,}6\, x10^{-20} cm^2$$

3) For a given light intensity I_0 (W/cm^2), the photon flux amounts to I_0/E, and the kinetic rate of uncaging event, $k_{uncaging}$, is given by [136]

$$k_{uncaging} = 2.3 \times \varepsilon_{uncaging} \times \frac{I_0}{E} \quad \textbf{(3)}$$

The light intensity I_0 (W/cm^2) of the UV lamp was measured using a powermeter. This calibration was done on the M365L2-C4 from Thor Labs which is a circular and power-tunable 365nm LED based lamp. Optics were added in the path of the light to ensure that the beam of light provides a uniform flux. I_0 was measured to be 45mW/cm^2. Using equation (3), the uncaging kinetic in this set up is found to be $k_{uncaging}$ = 0.0068 sec^{-1}.

The exponential kinetics of uncaging reads as

$$f(t) = 1 - e^{(-k_{uncaging} \times t)} \quad \textbf{(4)}$$

and reaches 63% for t= $\tau_{uncaging} = \frac{1}{k_{uncaging}}$ = 147 seconds

In order to reach approximately 100% of uncaging with less than 1% of remaining uncaged compound, the illumination duration used is such that

$$t = 5 \times \tau_{uncaging} \quad \textbf{(5)}$$

In other words, in these experimental conditions, the time required to uncage 63% of active cyclofen is about 2.4 minutes. This corresponds to a dose density or fluence of 6.48 J/cm^2, i.e. 6.48 µJ on a 100µm^2 surface, about the size of the surface area of a cell.

7.2.3 Experimental calibration of photoactivation using p-nitrophenol

To confirm the previous calibration, an experimental approach was performed. Given that caged cyclofen does not present any specific properties which would allow to measure the advancement of the uncaging reaction, this experimental calibration was done by absorption spectroscopy with a different caged substrate with similar uncaging kinetics. 4,5-dimetoxy-2-nitrobenzyl, or caged p-nitrophenol, has a similar caging group

as caged cyclofen, does not isomerize nor degrade under UV illumination, and present a strongly colored photoproduct that allows the study of its uncaging kinetics (Figure 7.1, D) [137]. A 25µM solution of caged p-nitrophenol in acetonitrile (ACN)/20mM Tris buffer pH 8.1 1:1 (v/v) solution was exposed for different durations to UV illumination at a power of 45mW/cm^2. The caged p-nitrophenol absorption band at 300nm continuously decreased as the exposure time increased, and an absorption band corresponding to the release of 4-nitrophenate anion increased at 395nm (Figure 7.2, a). The monoexponential reaction reaches a plateau phase when all the caged p-nitrophenol has been uncaged. Further illumination beyond this point will not release more anion. From the representation in (Figure 7.2, b), the uncaging time, $\tau_{uncaging}$, can be read as the time where 63% of the 4-nitrophenate anion has been released and thus when the 395nm absorbance band has reached 63% of its maximal value. From (Figure 7.2, b), the uncaging time is found to be $\tau_{uncaging} \approx 2$min. This value is in agreement with the theoretical value found previously suggesting that this experimental calibration is valid and can now be applied to release caged cyclofen in the biological model.

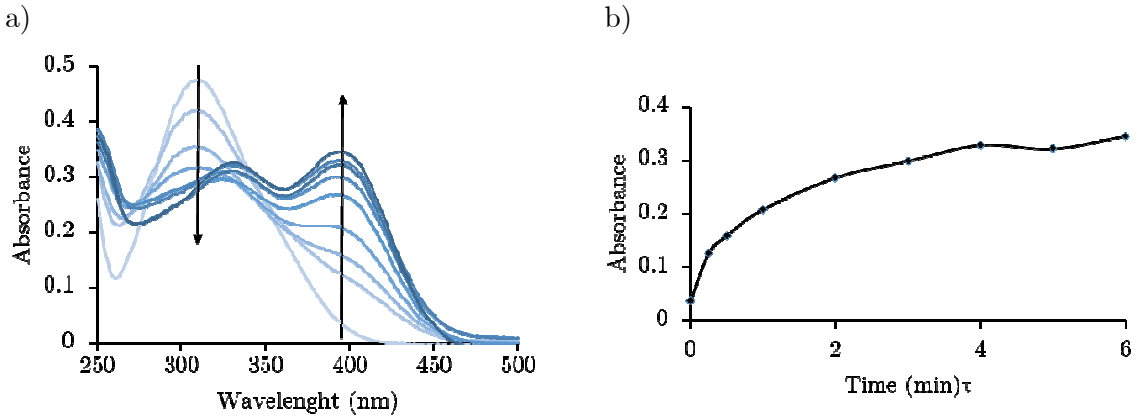


Figure 7.2 Kinetics of caged p-nitrophenol uncaging at 395nm

UV illumination (I_0=45mW/cm^2) of a 25µM solution of p-nitrophenol in acetonitrile/20mM Tris pH=8.1 buffer. a) Evolution of the absorption band shift from 300nm to 400nm with increasing exposure time (t(min)=0, 0.25, 0.5, 1, 2, 3, 4, 5, 6). b) Monoexponential uncaging reaction of caged p-nitrophenol at 395nm.

7.2.4 Homogeneous induction of EMT

In the previous sections, the experimental conditions for controlled photo release of cyclofen from its caged precursor have been established. Using this knowledge, it is now possible to set up an experiment where MDCK-SnailERT2 cells are incubated with the caged precursor, and then UV exposed to trigger an EMT response in a homogeneous manner (all the cells are induced). To limit the amount of UV light shown onto the cells

while ensuring that the photoactivation reaction is in the plateau phase and that sufficient cyclofen is being released, the duration of illumination was set to 5min. This value corresponds to about 2 x $\tau_{uncaging}$ that was previously calibrated. The response was assessed both by morphological images as well as through IF (Figure 7.3). In the negative control, when cells are incubated in DMSO and not subjected to UV illumination, morphological images reveal a tight and regular epithelium (Figure 7.3, A). This is further appreciated through intercellular staining of E-cadherin and ZO-1. Although the signal provided by the former has a lot of background noise, E-cadherin localization to the cell junctions is still evident and intercellular junctions integrity is validated with the clean ZO-1 staining. The mesenchymal staining for vimentin and F-actin present relatively low basal expression (Figure 7.3, 1st row). When cells are bathed in caged cyclofen and exposed to UV, cellular individuation and elongation appears (Figure 7.3, D). Apparition of spiky cellular edges and individualization of cells is seen morphologically. Stronger magnification in IF allows to appreciate the degree of cell spreading and dissociation. In this experimental condition, intercellular junctions are lost as seen through the reorganization of both E-cadherin and ZO-1 proteins. In addition, strong increase in the expression of vimentin, and apparition of stress fibers of F-actin reveal this important morphological change. Finally, the nuclear counterstaining in DAPI allows to appreciate the increased distance between the nuclei as a result of EMT induction (Figure 7.3, 4th row). Controls for the effect of caged cyclofen and UV illumination were done separately. In both cases, morphological images and IF analysis are comparable to the negative control situation (Figure 7.3, compare 2nd and 3rd with 1st row). This highlights two important facts. First, the presence of the caged precursor is not sufficient to induce EMT on its own, and specifically requires that UV illumination be added to see an effect. This proves that the system is not leaky and further validates the optogenetic model. Second, the UV illumination conditions used in these given experimental conditions are not lethal to the cells, and do not seem to affect them in any way. Possible toxicity effects caused by UV illumination were further studied. Proliferation assays were done to verify that the deposited dose of light did not modify neither cell division rate nor proliferation. FACS sorting was done on exposed cells to verify that DNA content in these cells was not modified (data not shown).

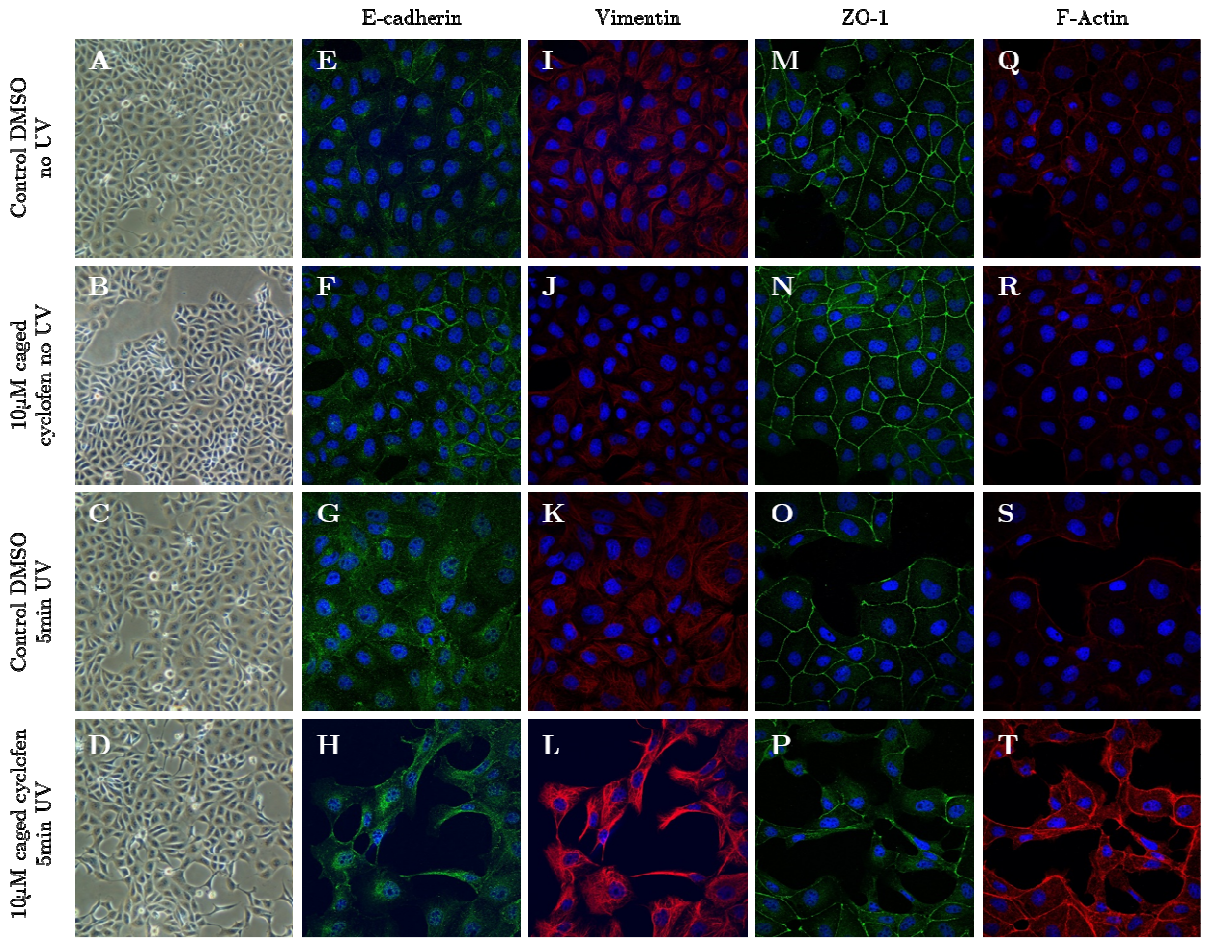


Figure 7.3 Homogeneous EMT induction in MDCK-SnailERT2 cells

Morphological images (A-D) and immunofluorescence stains for the indicated markers (E-T) in MDCK-SnailERT2. Cells were incubated for 1hr in starving media with 10μm caged cyclophen. Cells were then washed and illuminated for 5min to UV light, I_0/S=45mW/cm^2. After illumination, cells were washed again. Following a 24hr incubation, cells were fixed and stained for the adherent junction marker, E-Cadherin (E-H) as well as for the tight junction marker ZO-1 (M-P). Cytoskeletal organization was assessed by both Vimentin (I-L) and F-actin (Q-T) markings. Nuclei are counter-stained with DAPI (blue). (Detection channels: blue 405nm, green 488nm and red 635nm)

As previously, the EMT response was quantified using qPCR analysis. EMT induction with cyclofen causes a significant change in mRNA expression levels only in three of the five markers tested (E-cadherin, ZEB-1 and Slug) (Figure 6.4). As a result, the following experiments will focus only on these three markers. When cells are incubated in caged cyclofen and exposed to UV, transcriptional levels of E-cadherin show a significant decrease. In addition, ZEB-1 and Slug levels increase suggesting that photoactivation of EMT is taking place in a manner that is comparable to EMT induced

with constitutive cyclofen (Figure 7.4). Controls for the effect of UV illumination and caged cyclofen alone show no significant changes in the expression of these three EMT markers.

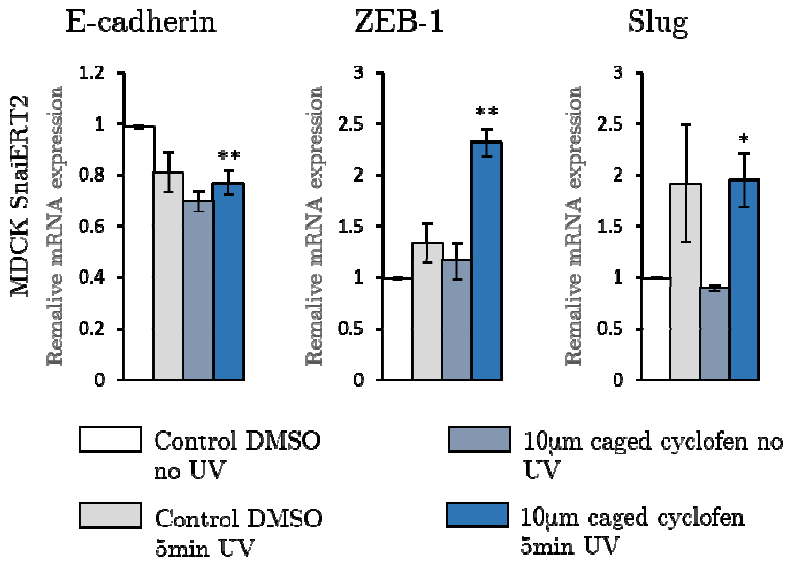


Figure 7.4 Homogeneous photoactivation of EMT in MDCK-SnailERT2 cells

Cells were incubated for 1hr in starving media with 10µm caged cyclophen. Cells were then washed and illuminated for 5min to UV light, I_0/S =45mW/cm^2. After illumination, cells were washed again. Following a 24hr incubation, cells were lysed for RNA extraction. qPCR analysis reveal significant changes in the expression of E-cadherin, ZEB-1 and Slug and E-cadherin in MDCK-SnailERT2. Error bars represent SEM, n=4, p≤0.05.

Results presented in this section show that caged cyclofen can be photoactivated in a non-toxic manner using UV illumination. In addition, the EMT response elicited optically is comparable to the response obtained with cyclofen. This confirms that the design and calibration of the system is correct, and that the model is suited to study situations of heterogeneous EMT.

7.3 Fluorescent reporter of the photoactivation event

7.3.1 Validation of the stable cell line MDCK-SnailERT2-H2B Dendra2

With the optogenetic approach now calibrated and functional, a last critical tool was needed to generate heterogeneous EMT. In fact, as mentioned in the introduction, it is hypothesized that in the situation of heterogeneous EMT induction, two possible outcomes may exist. The first one is *phenotypic propagation*, where the cell induced to

undergo EMT is sufficient to destabilize its neighbors through loss of cell junction stability. This would result in surrounding cells to undergo at least partial EMT although they did not receive the induction and their internal epithelial determinants were not directly perturbed. In this case, the observer would see epithelial dissociation and the effect would be macroscopic and measurable. The second possible outcome would be the reverse, where the cell induced to undergo EMT would in fact remain part of the epithelium through stabilizing actions of its neighboring cells. In the latter case, defined as *phenotypic suppression* by H. Rubin [138], the cell would have received the internal cue to undergo EMT, but would not be able to proceed through the morphological switch (Figure 7.5). To the observer, it then becomes problematic to assess whether a cell has been challenged or not, given that no effect is noticeable. To overcome this problem, and to be able to detect such situations of phenotypic suppression, a fluorescent reporter of the event of photoactivation has been developed. The green-to-red photoconvertible fluorescent protein Dendra2 [139, 140] was stably expressed in MDCK-SnailERT2 cells. The advantage of this approach is that photoconversion of Dendra2 takes place under the same UV illumination wavelength required for the photoactivation of cyclofen (Figure 7.6). Thus, in this model, the optogenetic approach will induce two distinct events: photo release of cyclofen from its caged precursor, as well as green-to-red photoconversion of Dendra2.

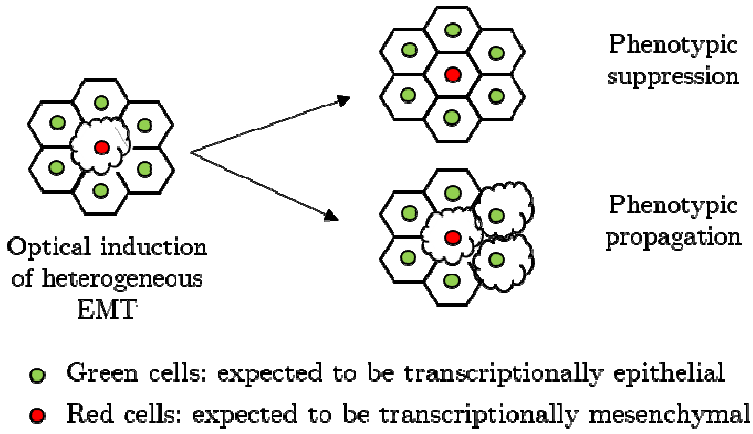


Figure 7.5 Possible outcomes of heterogeneous junctions

Using the H2B Dendra2 construct, it is now possible to keep a historical trace of which cell has been induced to undergo EMT. More specifically the GFP-like to RFP-like photoconversion will be a key tool to study the degree of cell-cell interplay and address the question of emerging properties.

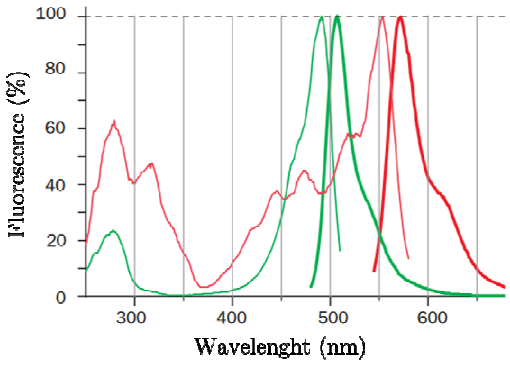


Figure 7.6 Fluorescence spectrum of Dendra2 protein

Normalized excitation (thin lines) and emission (thick lines) spectra for non-photoconverted (green lines) and photoconverted (red lines) Dendra2. (http://evrogen.com/products/Dendra2/Dendra2_Detailed_description.shtml.)

In the episomal plasmid used, the Dendra2 protein under the CMV promoter has been localized to the nucleus by fusing it to the C-terminal of histone H2B [101]. Localizing the fluorophore to the nucleus is important for imaging. If the Dendra2 protein was expressed in the cytosol, it would occupy two main imaging channels (GFP and RFP) and would limit the available channels to image membrane bound protein. This plasmid was further cloned in the pCDH1CMV-MCS-EF1-Puro lentiviral vector (System Biosciences) for increased infection yields and stable expression of the protein compared to a transient transfection. (Please refer to the Material and Method section for cloning strategy). This newly established cell line, referred to as MDCK-SnailERT2 H2B-Dendra2, was subjected to TGF-β and cyclofen treatment to ensure that the stable fluorophore did not interfere with the capacity of the cells to respond to the drug. In the same manner as Figure 6.1, cells were stimulated for 48hrs in 10ng/ml TGF-β or 24hrs of 5μm cyclofen (Figure 7.7). Samples were fixed and immunofluorescence staining of E-cadherin, F-actin, ZO-1 and Vimentin were done. Morphological images show that the cells react to both TGF-β and cyclofen in a similar manner than MDCK-SnailERT2 cells, showing cellular elongation and spiky edges respectively (compare Figure 7.7 with Figure 6.1). In addition, EMT in these cells induce a similar reorganization of both intercellular junction markers, E-cadherin and ZO-1. Stress fibers of F-actin and upregulation in the expression of Vimentin are noticed. Note that the nucleus of the cells appear, as expected, in the GFP detection channel (488nm). This is particularly true for the E-cadherin channel. The nuclear staining is less evident in the ZO-1 channel, and this can be explained by the fact that this EMT marker is located much more apically in the cells, such that the nucleus is out of focus when imaging ZO-1. Also, as expected, there is

no RFP-like signal in the detection channel (555nm) since the cells have not be exposed to UV light.

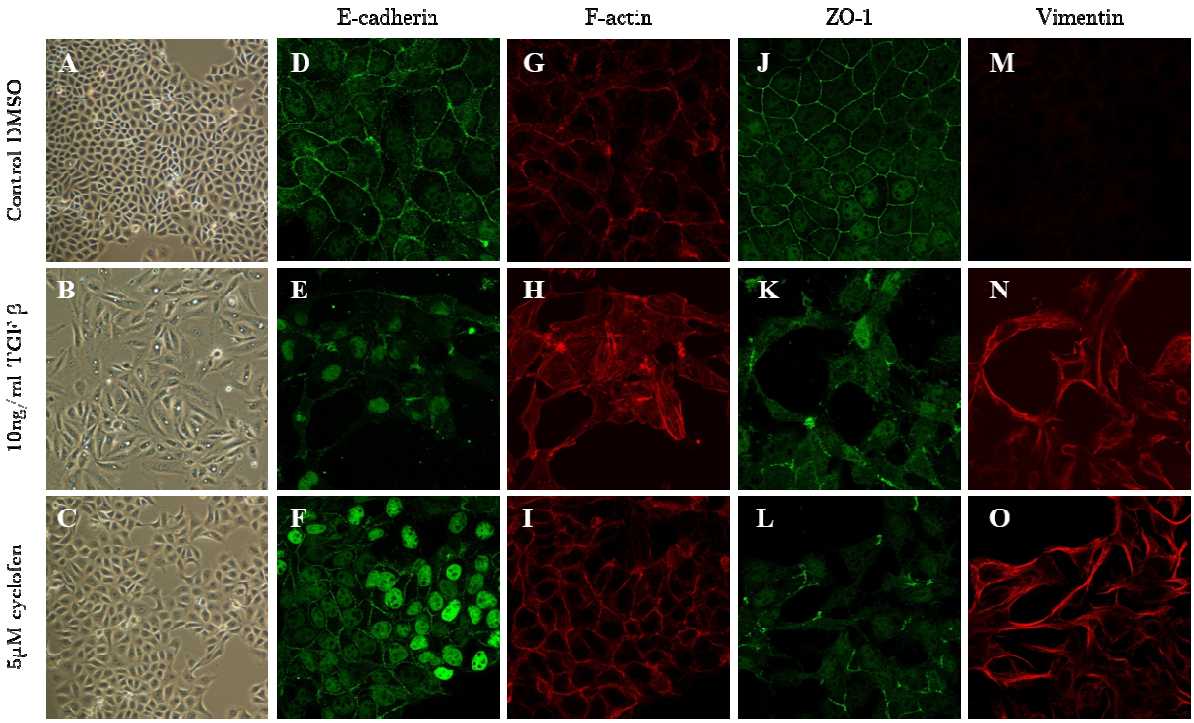


Figure 7.7 EMT induction in MDCK-SnailERT2 H2B Dendra2 cells

Morphological images (A-C) and immunofluorescence stains for the indicated markers (D-O) in MDCK Snail-ERT2 H2b Dendra2. Cells were imaged after 48hrs of DMSO (top row) or 10ng/ml TGF-β (middle row); cells were imaged after 24hrs following 20min incubation of 5μM cyclofen (bottom row). Cells were stained for the adherent junction marker, E-Cadherin (D-F) as well as for the tight junction marker ZO-1 (J-L). Cytoskeletal organization was assessed by both F-actin (G-I) and Vimentin (M-O) markings. (Detection channels: green 488nm and red 561nm)

7.3.2 Validation of photoconversion in photoactivated conditions

The H2B Dendra2 construct does not interfere with the established biological model as shown above. It is now necessary to verify that the protein is functional, and can undergo green-to-red photoconversion. Cells were exposed to UV illumination under the same settings established during the calibration of the uncaging of caged cyclofen. Fluorescence microscopy reveals that in these conditions, photoconversion of the H2B Dendra2 fluorophore takes place and can be assessed through the apparition of a RFP-like signal that is absent when cells are not exposed to UV light. As a result, it is thus possible to photoconvert the nucleus of illuminated cells to red under the experimental conditions required to uncage the active ligand of ERT2. This confirms the usability of

the H2B Dendra2 tool as a way of 'coloring' cells that have been subjected to UV illumination (Figure 7.8).

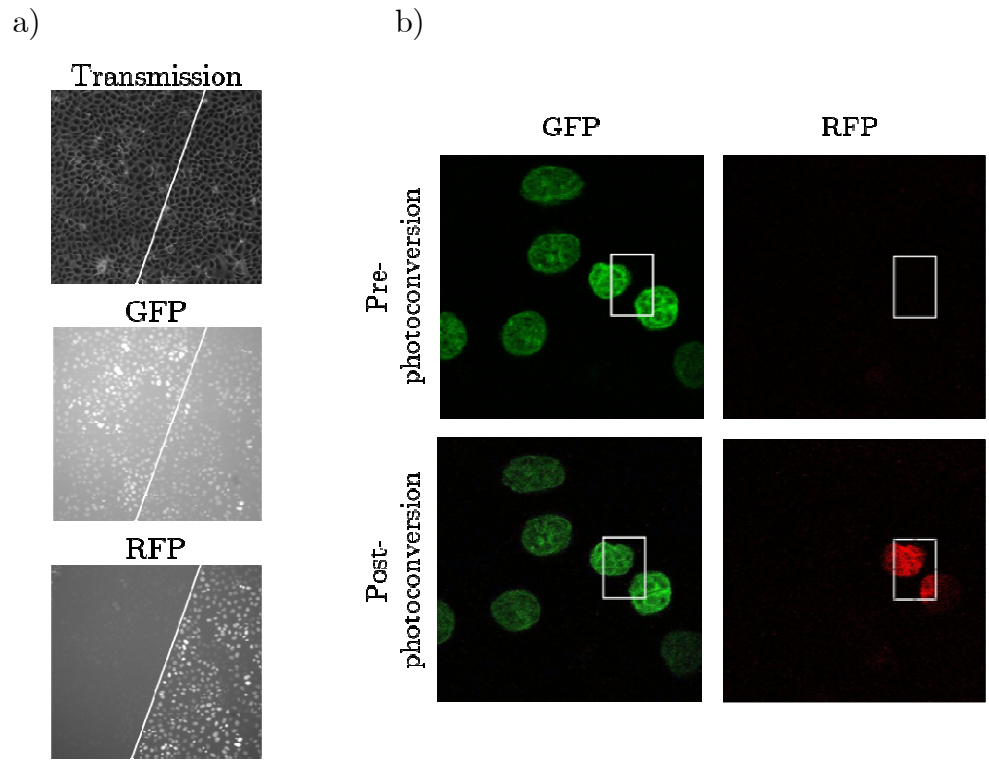


Figure 7.8 H2B Dendra2 green-to-red photoconversion in MDCK cells

MDCK-SnailERT2 H2B Dendra2 cells were exposed to UV illumination to assess the photoconversion event. a) Masked illumination using a 365nm UV lamp for 5min at I_0=45mW/cm^2 result in photoconversion only in cells exposed (white line marks the limit of the mask). b) Photoconversion was also done on a confocal microscope to appreciate the accuracy of photoconversion within a given nucleus (white square defines the region of 405nm laser illumination). (Detection channels: green 488nm and red 561nm)

The photoconversion event is irreversible but the ability to detect the RFP-like signal decreases over time. In fact, as cells divide, the fraction of photoconverted Dendra2 molecules will be split between the two daughter cells resulting in a weaker RFP-like signal in each individual cell. Experiments were carried out to measure the time limit beyond which it became challenging to detect photoconverted cells and is found to be about 30 hours (data not shown). This time window is in agreement with the previously established experimental frame set of 24hrs required to detect EMT.

7.3.3 Sorting photoconverted cells

With the MDCK-SnailERT2 H2B Dendra2 cell line, it is now possible to distinguish the cells that have been UV illuminated (and EMT induced in the presence of caged cyclophen) from those that have not been acted upon based on the color of their nucleus. In fact, cells with a GFP-like nucleus are expected, if cells are truly autonomous from their neighbors, to remain epithelial whereas those that have a RFP-like nucleus are expected to undergo EMT. For the remaining of this manuscript, and for ease in the reading, the un-induced cells expected to remain epithelial will be referred to as green cells, and those that are expected to undergo EMT as a result of cyclophen uncaging will be referred to as red cells (Figure 7.5), with reference to the color of their nucleus.

Using this colorimetric tool, it is now possible to use flow cytometry to FACS sort the two populations and isolate them from one another. The power of this approach is that transcriptomic analysis can be done on both populations separately to see if their transcriptomic profile corresponds to the expected epithelial or mesenchymal state. In other words, sorting the cells 24hrs after EMT has been induced and cell populations were let to interact, allows to study the degree to which cells are autonomous or not and address the question of emerging properties. If in fact the fate of a cell depends on the state of its neighbor, then this should be reflected by a mismatch between the expected transcriptomic state (dictated by the color of its nucleus) and its actual internal state (induced or not). In other words, if *phenotypic propagation* takes place, green cells that have not received the cue to undergo EMT will in fact have a less epithelial or more mesenchymal-like transcriptomic profile due to the signal received by the surrounding induced red cells. On the other hand, *phenotypic suppression* will lead to red cells presenting a less mesenchymal or more epithelial-like profile because of suppressive signaling received from the surrounding green cells (Figure 7.5). A typical workflow is presented in Figure 7.9.

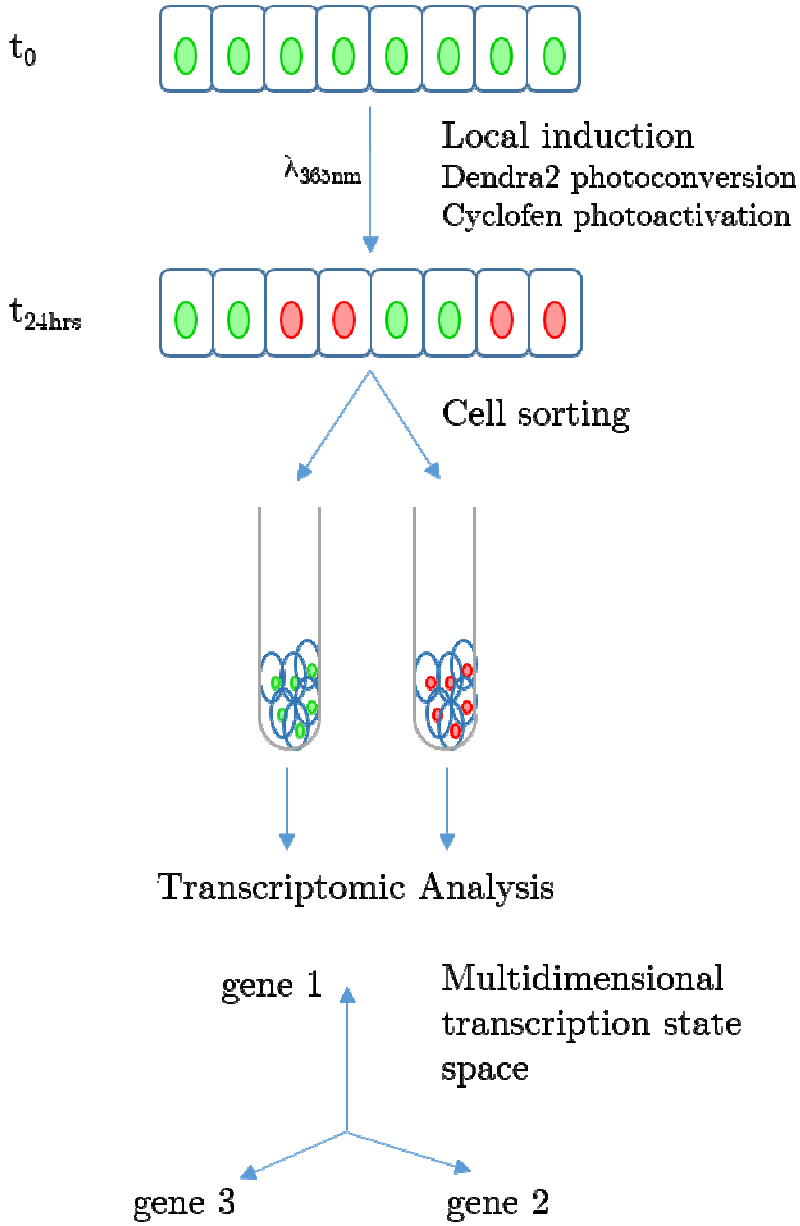
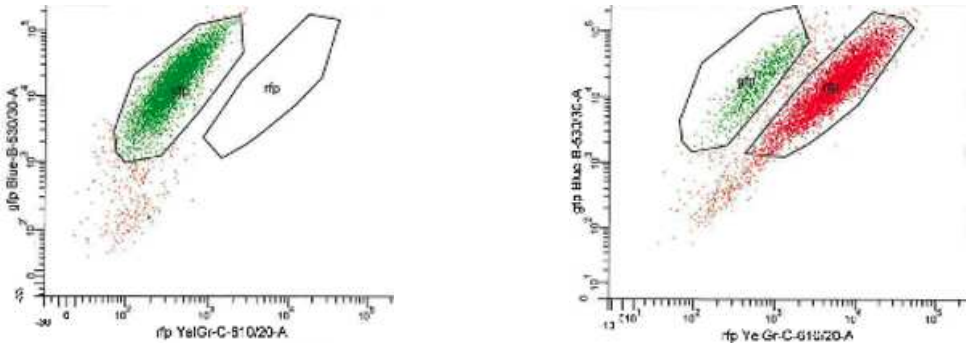


Figure 7.9 Experimental strategy

Cells are initially platted and grown to confluency. At t_0, the monolayer is incubated with caged cyclofen and local UV illumination is performed. The used wavelength allows for both, the uncaging of the active ligand, and photoconversion of the Dendra2 protein. The latter serves as a historical tracer to recover, after 24hrs of incubation, cells that have received the illumination from those that did not. Based on the nuclear staining, the two population can be sorted out and transcriptomic analysis can be performed on the two populations separately. Comparing the transcriptomic profile of these populations to that of homogeneously induced or homogeneously non-induced cells will bring out the effect of intercellular influences. These can be represented in a transcription state space.

Preliminary FACS sorting experiments were carried out to verify that the photoconversion event could in fact be a valid approach to precisely separate the two populations.

a) b)



No UV illumination Full dish UV illumination

Figure 7.10 FACS sorting of MDCK-SnailERT2 H2B Dendra2 cells

MDCK-SnailERT2 H2B Dendra2 cells were grown in 12 well dish and were either not illuminated (a) or exposed in a full dish configuration to UV illumination for 5 minutes with I_0=45mW/cm^2 (b). Cells were then trypsinized, filtered and sorted on a FACSAria (BD Biosciences) based on their GFP-like and RFP-like signal. Unstained and single colored cell lines were used to set the appropriate compensation and sorting gates.

From Figure 7.10, it is once again possible to confirm that there is no RFP-like signal prior to UV illumination (a) and that most of the cells present GFP staining. When all the cells of the dish are exposed to UV illumination, FACS analysis reveals the GFP-like to RFP-like photoconversion and allows to quite easily sort these two distinct populations. It is interesting to note that the apparition of the RFP-like signal is accompanied by almost no decrease in the expression of GFP-like. In other words, after photoconversion, cells are individually positive for both GFP and RFP. Surprisingly, when all the cells of the dish are subjected to UV illumination (Figure 7.10, b), a small portion of cells does not undergo photoconversion and remain in the GFP gate. The reason why the Dendra2 protein in these cell is not responding to UV is unknown, but suffice to say that it is important to consider this population for future experiments (Figure 7.11). Because photoconversion of Dendra2 and photoactivation of caged cyclofen are two independent events taking place in parallel, these refractory RFP negative cells will supposedly be induced for EMT. As a result, this lagging population is a source of false negative which is critical to take into account when analyzing the transcriptomic profile of the cells after EMT induction.

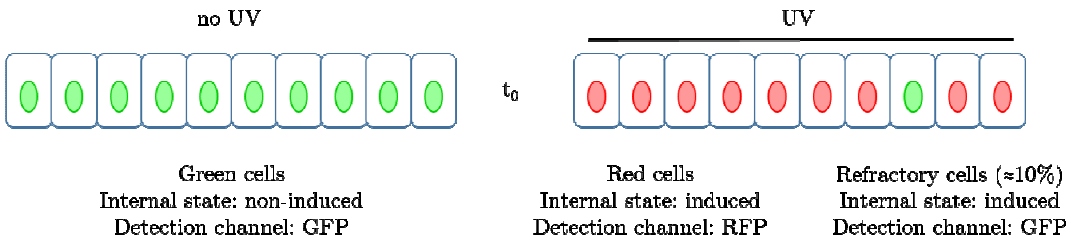


Figure 7.11 Refractory cells: a source of false negatives

Dendra2 undergoes green-to-red photoconversion when illuminated by UV light. A small population of illuminated cells will however remain RPF-like negative although they have received the necessary UV dose. These refractory cells represent about 10% of the illuminated population and are a potential source of false negatives.

Longer illumination durations have shown that this population of false negative does not undergo photoconversion, suggesting that the effect is not due to experimental conditions. Finally, this population represents about 10% of the total RFP positive population and the error it represents is included in the transcriptomic analysis.

7.4 Generation of heterogeneous illumination

In order to generate heterogeneous patterns of EMT, it is necessary that the light source be precisely deposited only on chosen cells. To generate this patterned illumination, three different approaches, with different advantages, have been developed and will be described in the following section.

7.4.1 Using illumination mask

The first approach was carried out using adhesive masks placed beneath the cell monolayer. This method allowed for the generation of macroscopic masks before the single cell illumination was available. Masks were done using black adhesive tape (Chatterton) that was cut out according to pre-defined patterns using a numerical cutter (Graphtec Craft Robo Pro). The capacity of this material to block UV was tested by placing a piece of tape in a spectrometer and measuring the light passing through in the UV range. The UV impermeable nature of the tape is important to ensure that only cells exposed (not masked) are being illuminated, such that only in these cells should EMT take place. Resolution of the mask was measured to be about 25µm which corresponds to more or less the size of a MDCK cell.

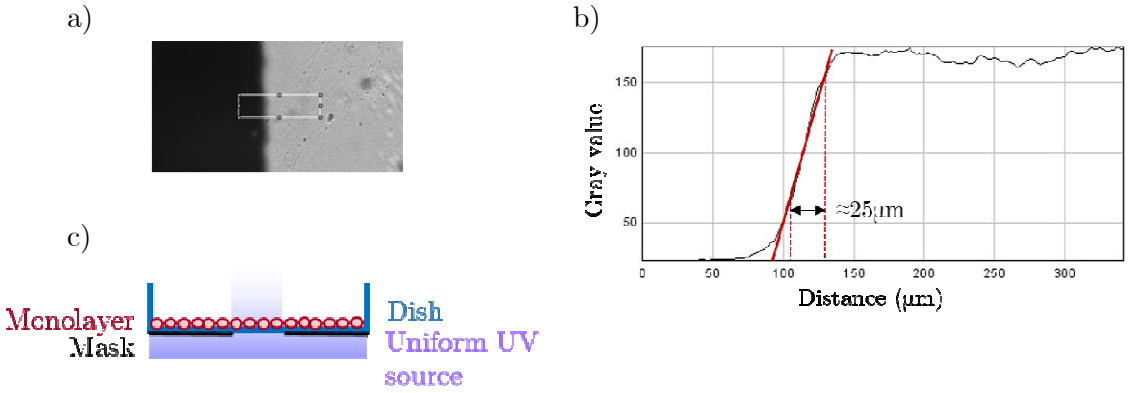


Figure 7.12 Characterization of Chatterton mask

Measure of mask resolution done on an image taken at the border of the tape (white rectangle) (a). The slope of the change in intensity gives a value of the resolution as the typical range found between 10 and 90% of the maximal contrast in the linear portion of the curve (b). The adhesive mask is placed beneath the cell culture dish, and cells are illuminated from the bottom using a uniform UV light source (c). Only cells not protected by the mask will be illuminated.

The advantage of this approach is the ease with which the masks are done. In addition, this method does not require any specific optical setup to generate patterned illumination a part from a collimated light source. A drawback is that patterning is not unlimited in that a new mask needs to be generated for each new pattern. In addition, this approach only allows for macroscopic scale patterns due to the limited resolution of the numerical cutter of about 250µm. Most of the data presented in this work has been generated with this method, but parallel work was focused on setting up alternatives to reach the single cell level.

7.4.2 Using holographic illumination

The second approach used to generate patterned illumination allows to reach the single cell level. In this project, a Liquid Crystal Spatial Light Modulator (LC-SLM, Holoeye) was used to modulate the phase of the incident beam [141]; the amplitude of the incident light source is not affected. A spatial light modulator is made up of an array of individual pixels, each of which controls its electrical potential independently. These pixels, controlled by an electric field, cause the alignment of overlaying liquid crystal molecules. The latter will undergo a tilt according to the electric field they are subjected to, causing birefringence and a change in the optical path length in the liquid crystal layer. This leads to phase modulation of the light as a function of its polarization. Holographic illumination imposes spatial modulation on a beam of light, allowing precise and flexible light excitation patterns with relatively high diffraction efficiency [142]. In

addition, this approach allows for the synchronous simulation of several cells due to the large excitation area.

Holographic illumination consists in calculating a phase pattern that creates the desired pattern to be illuminated at the objective focal plane. This is done using an iterative algorithm provided with the LC-SLM. The computer generated hologram (CGH) is then projected onto the LC-SLM that has been placed in the path of the excitation source such that it is conjugated with the back focal plane of the objective. The beam is then focused onto the sample plane, reproducing the desired template (Figure 7.13, a and b).

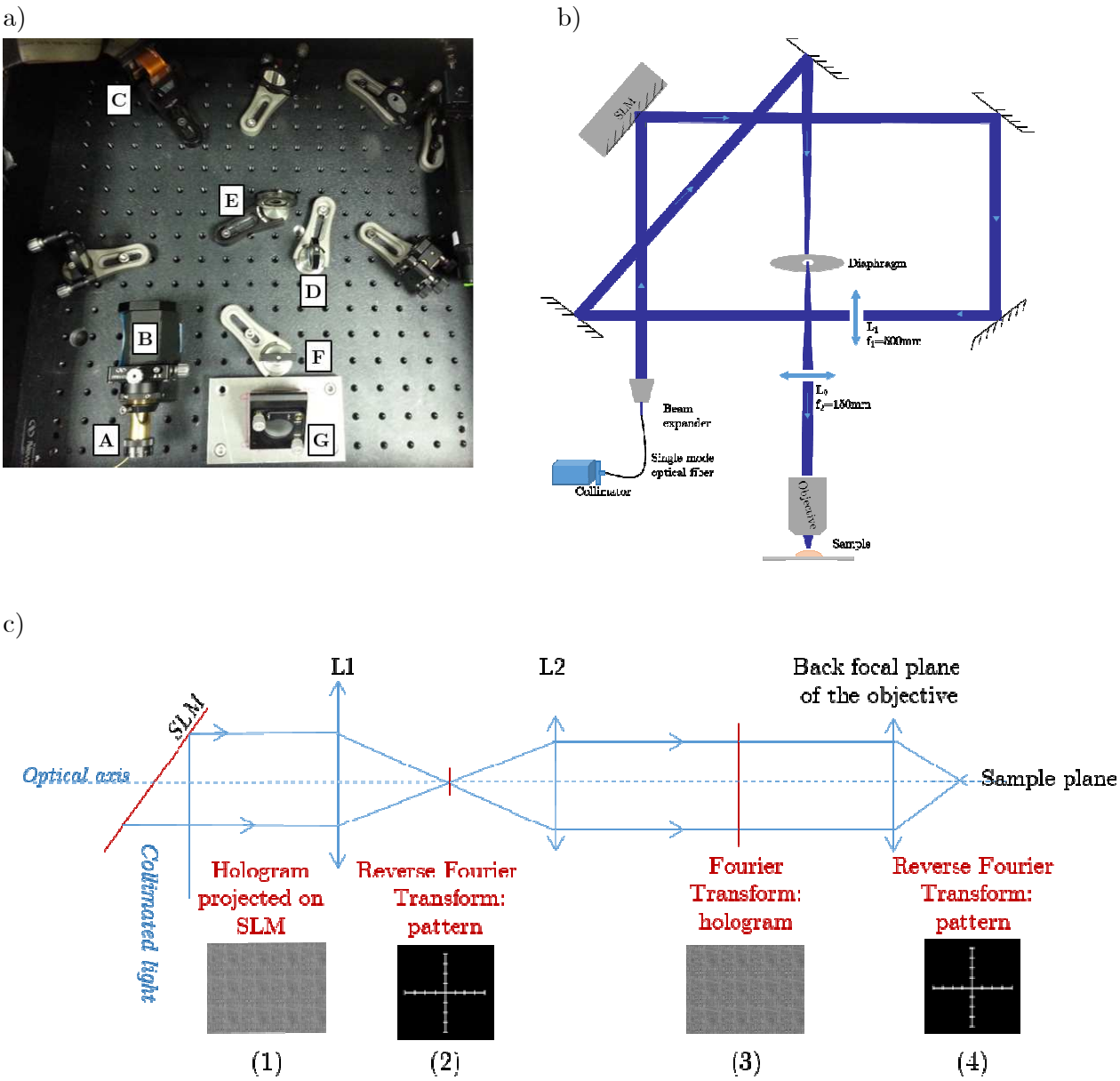


Figure 7.13 Optical setup for holographic illumination

Layout of the optical setup for the use of a LC-SLM. a) Experimental setup built for holographic illumination. Components are as follow: A: Polarization-maintaining single mode optical fiber; B: Beam expander; C: LC-SLM; D: Lens L_1 (f_1=500mm); E: Diaphragm; F: Lens L_2 (f_2=150mm); G: Back focal plane of the objective. b) Schematic representation of the path of light in this setup. c) Alternating Fourier transformation to illuminate the sample plane with the desired patterned illumination.

In mathematical terms, the CGH is a Fourier transform of the desired pattern of illumination. The beam reflected from the SLM is then Fourier transformed on a Fourier plane by a 500mm focal length lens L_1 thus forming a first image of the template at the level of the diaphragm, placed at the focal plane of L_1. This same diaphragm is also placed at the focal plane of L2, a 150mm focal length lens, which is conjugated with the back focal plane of the objective. L_2 causes a reverse Fourier transformation of the beam modulation, the equivalent of the initial hologram. As light continues its path to reach the sample plane, a final Fourier transform takes place and results in in-focus sample illumination according to the initial template used to generate the CGH (Figure 7.13, c).

Holographic illumination presents the advantage of infinite and almost instantaneous patterned illumination possibilities. In addition, with this approach, most of the laser beam is modulated such that the total power is distributed within the desired motif of illumination with some intensity loss. SLM are very powerful at concentrating the laser power to illuminate small regions. This is not the case in other dynamic patterned illumination methods such as digital mirror devices (DMD), where undesired light is blocked out of the excitation field, resulting in a loss of a large fraction of laser power [143]. That said, there are nonetheless several important points to address prior to using a functional LC-SLM setup. The first optimization done was to replace the optical fiber by a polarization maintaining optical fiber. Indeed, because the LC-SLM used is modulating the phase only, it is important to have a linearly, non-changing, polarized incident beam. In this condition, the polarization can be aligned with the direction of the liquid crystal molecules. When the incident beam is not linearly polarized, the phase component that is not aligned with the liquid crystal molecules will be improperly modulated. Experimentally, the polarization maintaining optical fiber was mounted on the setup. The ideal orientation of the optical fiber was selected to generate the best modulated signal at the sample plane. This significantly improved the resolution of the pattern of illumination at the sample plane compared to when a non-polarization maintaining optical fiber is used. In the latter case only a single phase of light is properly being modulated, resulting in diffraction of the unmodulated light, and a decrease in contrast and resolution.

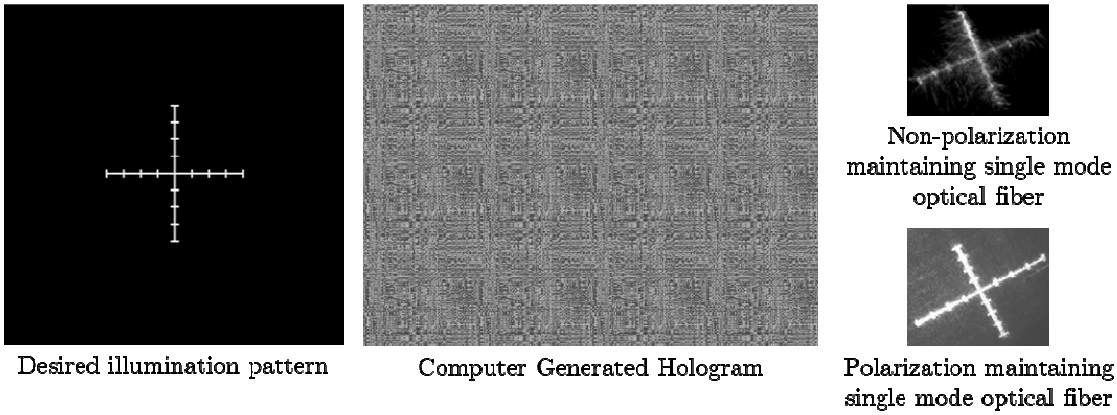


Figure 7.14 Effect of polarization maintaining optical fiber

Comparison of the reconstructed illumination pattern with a non-polarization maintaining monomode optical fiber and with a polarization maintaining optical fiber.

The second important point concerns the zero-order diffraction (ZOD) and how it is critical to remove it from the optical axis. The main origin of the ZOD from a phase only LC-SLM comes from the structure of the device itself. The LC-SLM is made up of an array of individual pixels which are separated by a minimal distance referred to as inactive region or dead space [144, 145]. Light arriving in these regions will not be modulated, resulting in a typical diffraction motif and generating a bright spot on the optical axis at the Fourier plane. Light power in the ZOD is in part dependent on the specific displayed pattern and is usually important, resulting in aberrantly high intensities in the center of the Fourier plane. Several approaches have been proposed to remove this undesired component. A crude approach is to introduce a beam block in the intermediate image to physically block the unmodulated light of the ZOD. Although efficient, this method introduces a blind region where modulated light that would localize to the same region as the ZOD will also be blocked out [142]. The second approach is to axially displace the ZOD away from the objective focal plane. In this case, it is necessary to slightly un-collimate the incident beam. This will allow the 1st order diffraction to remain on the objective focal plane while shifting the ZOD away (with a slightly convergent beam) or beyond (with a slightly divergent beam) the objective focal plane [141]. With this axial shift, it is now possible to include a beam block to remove only the ZOD without affecting the 1st order diffraction. It is important to state that this method can contribute to a decrease in diffraction efficiency [142]. A final approach to overcome this issue is to modify the hologram phase pattern so that it is shifted laterally from the center of the SLM, and thus of the ZOD. On top of this, it is also possible to apply to the hologram an additional spot which would generate destructive interference with the

zero order diffraction. This, however, requires precise knowledge of the phase of the zero order beam which is not an intrinsic characteristic but rather depends on the phase pattern itself [146].

Beyond these general considerations, and as the setup was being built and optimized it appeared more obvious that this method was not likely the most robust nor user friendly to use for single cell EMT induction. In fact, before any experiments were carried out, it was realized that the calibration of the laser power in each individual illuminated spot of the reconstructed pattern would be very difficult to accomplish. It was thus decided to use another method for single cell illumination where the same beam of light would be used to illuminate the different cells. In this way, the synchronous illumination is lost, but this allows to be certain that the dose of light deposited on a given cell will be similar to the dose deposited on the next cell as the beam of light remains unchanged. The tight control of the deposited dose is crucial to be able to affirm that the effect one may see in different cells results from the same identical initial perturbation.

7.4.3 Using neighboring algorithm

The goal in this approach is not only to locally induce EMT in individual cells as the SLM method would have allowed, but also to specifically choose the cells to illuminate according to strict spatial rules, or motifs. For this specific need, an algorithm was developed to automatically generate a graph of neighbors, and select the cells to be illuminated based on these neighborhood relationships. This allows to generate automated patterns with varying parameters (Figure 7.15). The entire workflow is based on confocal microscopy and in-house developed software in collaboration with Thomas Walter (Curie Institute, Mines ParisTech). *Selective Illumination* is an open-source software that was specifically designed for this study.

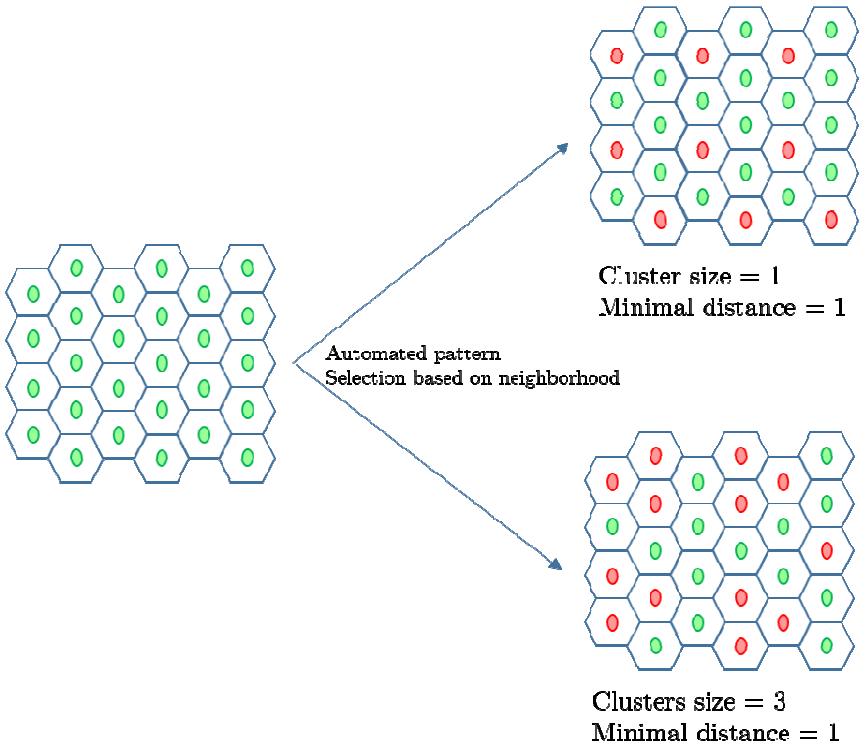


Figure 7.15 Automated generation of neighborhood patterns

The automated selection of cells based on their neighborhood would allow to induce EMT in individual cells according to well defined spatial rules. Varying the geometry of the distribution, or motif, of EMT induction would hypothetically allow to reveal the degree of cellular interdependence. Parameters such as cluster size, or minimal distance to the next induced cell can be varied, and the outcome of these situations can be compared.

The algorithm generates a list of coordinates that corresponds to the nucleus of cells to be illuminated according to strict spatial rules. The functioning of this algorithm is as follow. An image of the cell nucleus is done using the GFP-like fluorescence of the MDCK-SnailERT2-H2B-Dendra2 cell line (Figure 7.16). The image is then segmented to remove background noise, nuclear debris, and isolate the labelled individual cell nuclei. Cells and small objects touching the border of the field of view are removed to ensure that a cell is not partially illuminated. From the segmented nuclei, an area is constructed and expanded until it touches another surrounding area, or a certain distance from the nucleus is reached. This amounts to a Voronoï tessellation with an additional distance constraint (Figure 7.16, b). Touching regions are marked as adjacent, allowing for a precise determination of direct neighbors and the generation of a neighbor graph. From this network, a first cell is then chosen to be illuminated: either randomly or to best match the 'average' cell in terms of intensity, shape, texture (Figure 7.16, c). A more thorough description of these mathematical steps can be found in the Appendix.

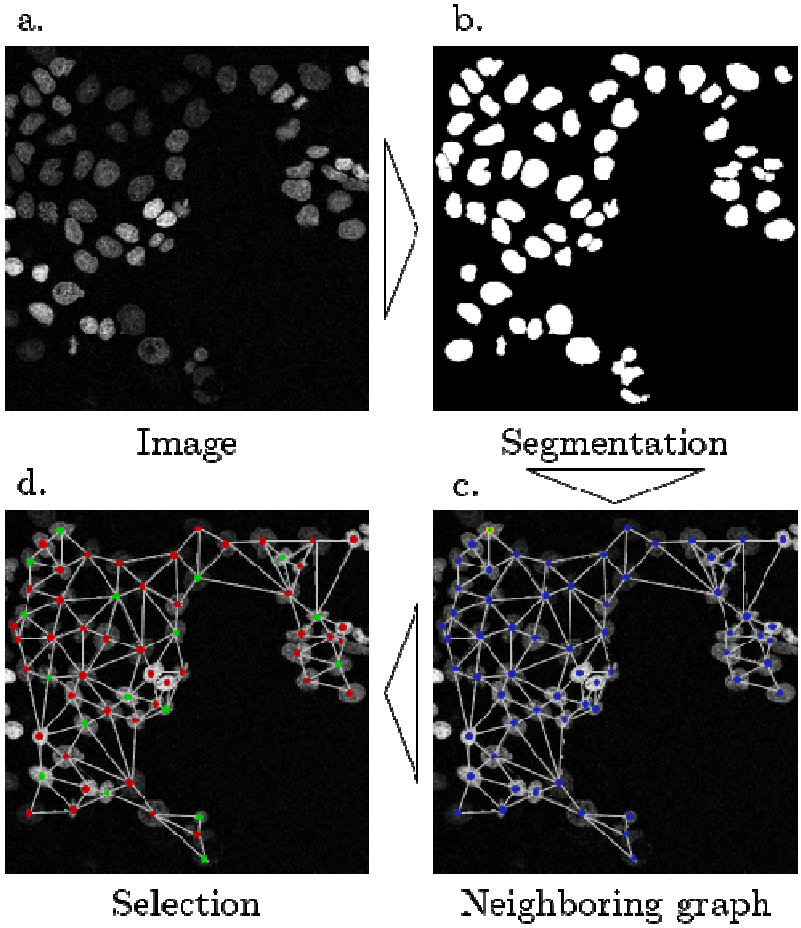


Figure 7.16 Generation of patterns for phenotypic induction

(a) Image of GFP-like signal in MDCK SnailERT2-H2B-Dendra2 cells, (b) Nuclei segmentation and neighborhood Voronoï tessellation, (c) neighborhood graph and a choosing of a most representative cell (blue: not processed cells; green with yellow circle: first chosen candidate cell to be illuminated), (d) final pattern generated by the software: selected cells (green) and not selected cells (red).

According to the spatial rule of illumination, neighboring cells are either considered for illumination or forbidden (Figure 7.16, d). Parameters such as minimal distance to the next considered cell, or size of the illuminated cell clusters can be defined: one every second cell, one every 10 cells, or clusters of three cells are typical spatial rules that can be implemented (Figure 7.15). These geometric features are designed to bring out the possible existence of a critical size for clusters to nucleate EMT or of a critical distance between cells. By iterating this algorithm to the next considered cell, the whole field of view can be covered, and a map of the nucleus to be illuminated will be generated (Figure 7.16, d). The position of each of the chosen cells is recorded and an

offset is applied to sequentially place each position at the center of the image. This ensures that during the illumination, the sample rather than the laser beam is moving thus controlling that the illumination dose remains unchanged from one cell to the next one. A list of position is thus generated and sent back to the confocal microscope. The illumination step is done through a macro which orders the automated stage to move such that the first position of the list is placed at the center. Once illumination has been done on this position for the defined duration, the stage places the second position, and this is repeated to all the positions of the list.

This method takes into consideration direct interacting neighbors, and specifically illuminates individual cells. Because it is at the single cell level, this method will allow to have more control on the motif of induction, and thus should enable a better understanding of potential rules and emerging properties.

This selection algorithm creates a list of cells that could in principle be illuminated simultaneously by the SLM to synchronize the photoactivation. Given that the time frame of the experiments is long (24hrs) compared to the time needed to distribute the illumination, the synchronization is in fact not an issue here. The major problem with the SLM however is that it would be very difficult to ensure that cells receive the exact same light dose, and to quantify this dose. In *Selective Illumination*, the calibration is made more easily by positioning the cell to illuminate at the center of the image thus ensuring that the same beam, and so dose, is used for all the cells.

Optimal photoactivation would entail that the whole cell must be illuminated in order to uncage all the molecules. However, because caged cyclofen is freely diffusing in the cytosol and hopefully also in the nucleus, it is possible to uncaged all molecules by illuminating only a fraction of the cell during a long enough time. This duration must allow all the caged molecules in the cell to diffuse to the region of illumination during this given period of time. In this method, the illumination region of a cell has been set to match to the average surface of the nucleus. This prevents from depositing light to neighboring non-selected cells. In addition, it ensures not only effective GFP to RFP photoconversion of the Dendra2, but also that no light is deposited without being reported by Dendra2 photoconversion. Because the size of the laser beam is much smaller than the surface of a typical nucleus, illumination is done by scanning this surface. As a result, to estimate the duration of illumination required to fully uncage all the molecules of cyclofen, a complex combination of diffusing molecules and a moving illumination volume needs to be considered.

When scanning the projected surface of the nucleus with the photoactivation point spread function (PSF), a series of successive sources of activated molecules are created from which diffusion occurs. Among these sources, those that are close to the

boundary of the nucleus will bridge the gap between the nucleus and the cell periphery faster than those in the middle of the nucleus. Photoactivation illumination done by scanning is thus more efficient than keeping the illumination PSF immobile at the center of the nucleus.

The time required for sources at the periphery of the nucleus to fill the gap by diffusion can be estimated by,

$$\tau_{diffusion} = \frac{\sigma_{gap}^2 - \sigma_{PSF}^2}{2D} \quad (4)$$

Where $\sigma_{gap} \approx 3\mu m$ is the typical distance between the border of the nucleus and the cell periphery (Figure 7.17), D = $4\mu m^2/s$ is the diffusion coefficient of the caged cyclofen [103], and σ_{PSF} corresponds to the radius of the activation point spread function (PSF). The latter can be calculated using the imaging setup specifications. The uncaging was performed with a 405nm laser using a 20x objective with a numerical aperture of NA=0.75, so:

$$\sigma_{PSF} = \frac{0.61\,\lambda}{NA} = 329\,nm = 0{,}33\mu m \quad (4)$$

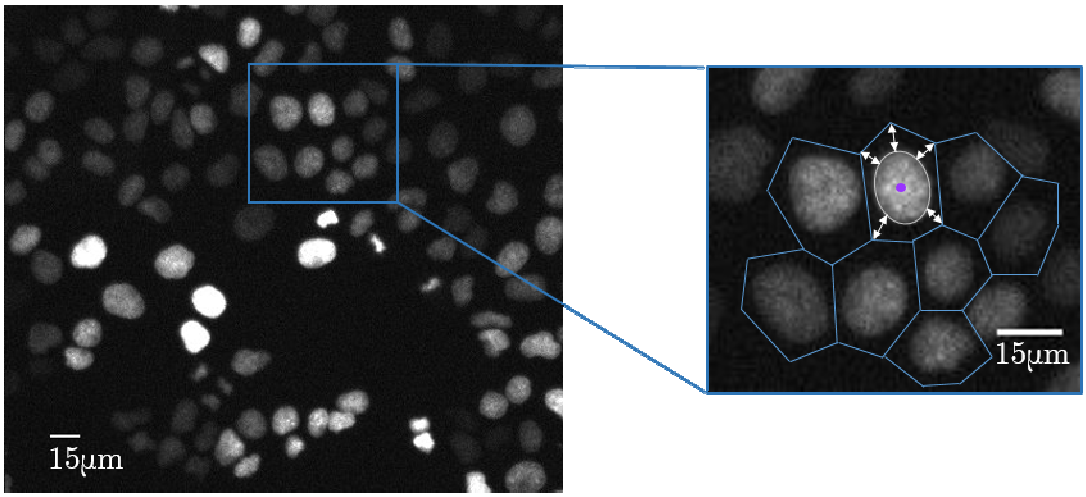


Figure 7.17 Diffusion front of a typical MDCK cell

Images represent GFP-like staining from MDCK SnailERT2-H2B-Dendra2 cells taken on a confocal microscope. Possible intercellular junctions are presented in blown-up image. White double-headed arrows show the gap to fill by diffusion of active cyclofen and is measured to be about 3-4µm wide. Purple spot represent (not to scale) the UV laser illumination.

How the diffusion front spreads over time can be plotted and is presented by its coordinate σ (σ=0 at the source center) in Figure 7.18.

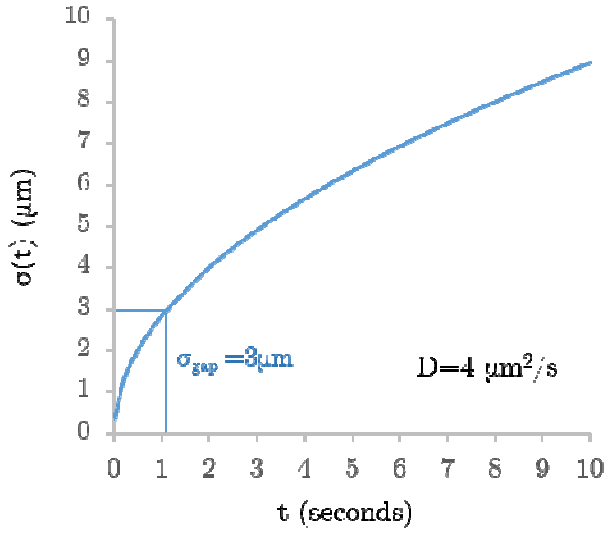


Figure 7.18 Representation of σ over time

Spreading of the diffusion front of caged cyclofen as a function of time. The gap between the boundary of the nucleus and the cell membrane is about 3 μm. The time needed for the molecules to travel such distance is about 1 second.

From the graph in Figure 7.18, the front of the distribution of activated molecules spreads from the boundary of the nucleus by 3-4μm in about 1-2 second. Based on this calculation, the experiments were carried out using 1 second of illumination time per cell. Importantly, it is of paramount importance that all cells receive the exact same amount of light. This is achieved by scanning a standardized disk, but there is no certainty as whether this procedure is optimal. Indeed, the bulk of the excited volume contains the nucleus, and the nuclear concentration of cyclofen is not known.

8 Heterogeneous induction of EMT

At this point, all the necessary tools to generate heterogeneous EMT patterns in a stable and homogeneous epithelium have been calibrated and established. The next section will present the most significant results obtained with the optogenetic approach. The first few experiments presented relied on the illumination mask to generate the illumination patterns. The last experiment was done using the *Selective Illumination* algorithm based method.

8.1.1 Radial distribution of EMT

In this first experiment of heterogeneous induction of EMT, no prior knowledge was known as to which potential parameters could have the strongest role in determining

the outcome in heterogeneous conditions. To question whether a specific spatial distribution of EMT induction or a critical distance with the next EMT induced cells are important parameters, it was necessary to generate patterns where both green and red cell populations are equal in proportion. In fact, this will ease the understanding of potential emerging properties by ruling out population size effects. Because the experiments were done in 12 well plates, a very natural pattern was a pie-like motif. The latter was generated with the Chatterton masks, and the experimental conditions consisted in differentially fragmenting the pie: one in two, one in four and one in eight. Schematically, the conditions of this experiment are represented in Figure 8.1 below the graphs. Briefly, disks represent the cell culture dish, and the colors represent the fluorescence of the nucleus in each portions of the dish. In other words, the distribution of the green cells describe the border of the Chatterton mask used. This type of representation will be used for the following experiments. As previously mentioned, green represents the cells that have not been illuminated and should be transcriptionally epithelial whereas red correspond to the illuminated cells that have been induced to undergo EMT and whose nucleus have undergone photoconversion. All the conditions represented have been incubated in caged cyclofen except the negative control (represented as a green disk) which was bathed in DMSO. The control for the effect of caged cyclofen alone and UV illumination alone were also done but are not presented here as their expression levels were comparable to the negative control. For practical reasons, the UV illumination was set to 28mW/cm^2. In these conditions, the illumination the $\tau_{uncaging}$ = 231sec = 3,8 min according to Equation (5). To ensure being in the plateau phase of the uncaging reaction, the experimental value used throughout the following experiments was set to $\tau_{uncaging}$ = 7min. This corresponds to a fluence value of 6,468 µJ for 100µm^2, a value comparable to the one found in Section 7.2.2.

Cells were incubated for 24hrs after cyclofen uncaging and the monolayers were sorted based on fluorescence. Expected ratios of 1:1 of the two populations were recovered during sorting. Transcriptional analysis was done on green and red cell populations similar to the homogeneous induction experiment (Figure 7.4), and the mRNA expression levels of E-cadherin, ZEB-1 and Slug were studied by qPCR.

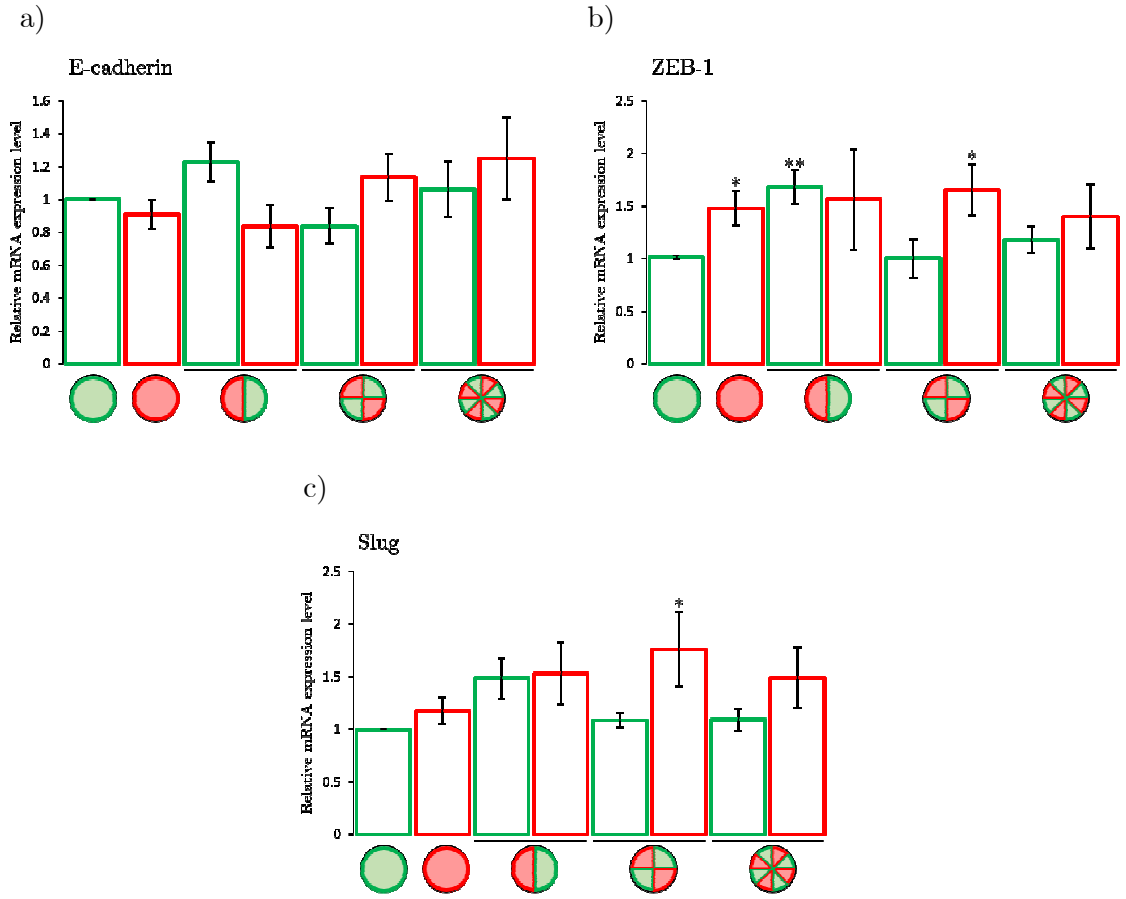


Figure 8.1 mRNA expression levels of EMT markers following radial distribution of EMT

MDCK-SnailERT2 H2B Dendra2 cells were either incubated in DMSO (negative control, or green disk) or in 10μm caged cyclofen for 1hr then washed (positive control and experimental conditions). The monolayer was illuminated for 7min at I_0=28mW/cm^2 with a 365nm UV light source and cells were washed again. Cells were incubated for 24hrs and then FACS sorted. mRNA from the green and red cell population was purified and qPCR analysis were performed on E-cadherin, ZEB-1 and Slug. Color code: green represents the green cells, or those that have not been illuminated. Red represents the red cells, or those that have been illuminated. Error bars represents SEM, n=5, p≤0.05.

Surprisingly, both positive controls for E-cadherin and Slug are not varying significantly such that is not possible to conclude on the expression levels of these genes. On the other hand, the response of the cells to ZEB-1 expression show striking results. In fact, if green and red cells behaved completely independent one from the other, the response to illumination should be comparable to the homogeneous induction (Figure 7.4). This is the case in the condition where the dish is fractioned in four: the green cells are not different from the negative control, and the red cells present an increased mRNA

expression level. However, the response is quite different when the dish is fractioned in two and in eight. In the former case, the green cells show significantly increased levels of ZEB-1 mRNA while the red cells do not vary significantly. This observation can only be explained by the fact that the two populations are co-existing and interacting. In fact, it appears as though the sole presence of the red cells is triggering this increase in ZEB-1 levels in the green cells. In addition, the effect also seems to be bidirectional as the red cells do not present the characteristic increased ZEB-1 levels of mesenchymal. In other words it is as if the green cells are preventing the red cells from becoming mesenchymal while the latter are inducing the green cells to be less epithelial-like. When the dish is fractioned in eight, the red cells appear more epithelial-like than they are expected to be. The data can also be plotted in 3D for a more complete visualization of the changes induced by heterogeneous EMT induction (Figure 8.2). In this representation, levels of all three genes can be included and it appears clear that the green cells of the dish portioned in two are in a region in space that is significantly different from the negative control and somewhat comparable to the region in space occupied by red cells. It is important to state that in this experiment and in the next heterogeneous experiments to be described, the morphology of the cells was not sufficient to appreciate nucleation or frustration events, explaining why the analysis is purely based in the transcriptomic data.

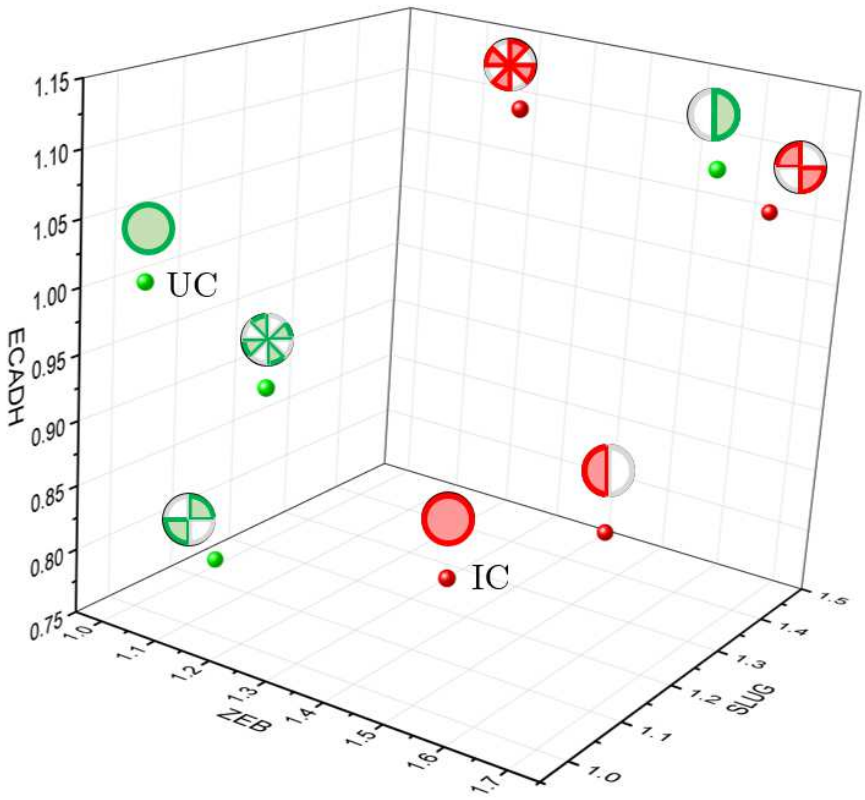


Figure 8.2 3D representation of mRNA expression levels of radially distributed EMT

E-cadherin, ZEB-1 and Slug mRNA expression levels were represented on a single 3D graph. This representation brings out the existence of emerging properties. (UC: homogeneously un-induced control, IC: homogeneously induced control)

The data obtained from this experiment suggest the existence of emerging properties that can only be explained because induced and non-induced cells are co-existing and forming a heterogeneous ensemble. However, it is difficult to conclude further as this type of patterning is problematic. In fact, the activation pattern is such that a cell at the center of the population will have a very different neighboring landscape than a cell that is at the border between the two populations. The ideal motif to rule out this issue would be an infinite checkboard where all the cells of a given population have the same distribution of interdistances with all activated and inactivated cells. This issue is important and has led to the generation of more periodic patterns to limit the contribution of these varying interdistances.

8.1.2 Equal sized band patterns

Following this first experiment, the next pattern that was generated were band motifs. In this configuration, the issue of intra population heterogeneity is limited due to

the periodicity in at least one dimension. The experiment was carried out similarly to the radial distribution of EMT. In this case however the motifs were bands of varying width: 0.5mm, 1mm and 5mm. The band motifs are schematically represented below the histograms. The negative control is represented as a straight line whereas the positive control is an upside down U shaped symbol. The color of the histogram band reflects whether cells have been illuminated (red) or not (green). Again, the 1:1 ratio of population size was maintained during illumination and were recovered during FACS sorting. mRNA expression levels of E-cadherin, ZEB-1 and Slug were analyzed.

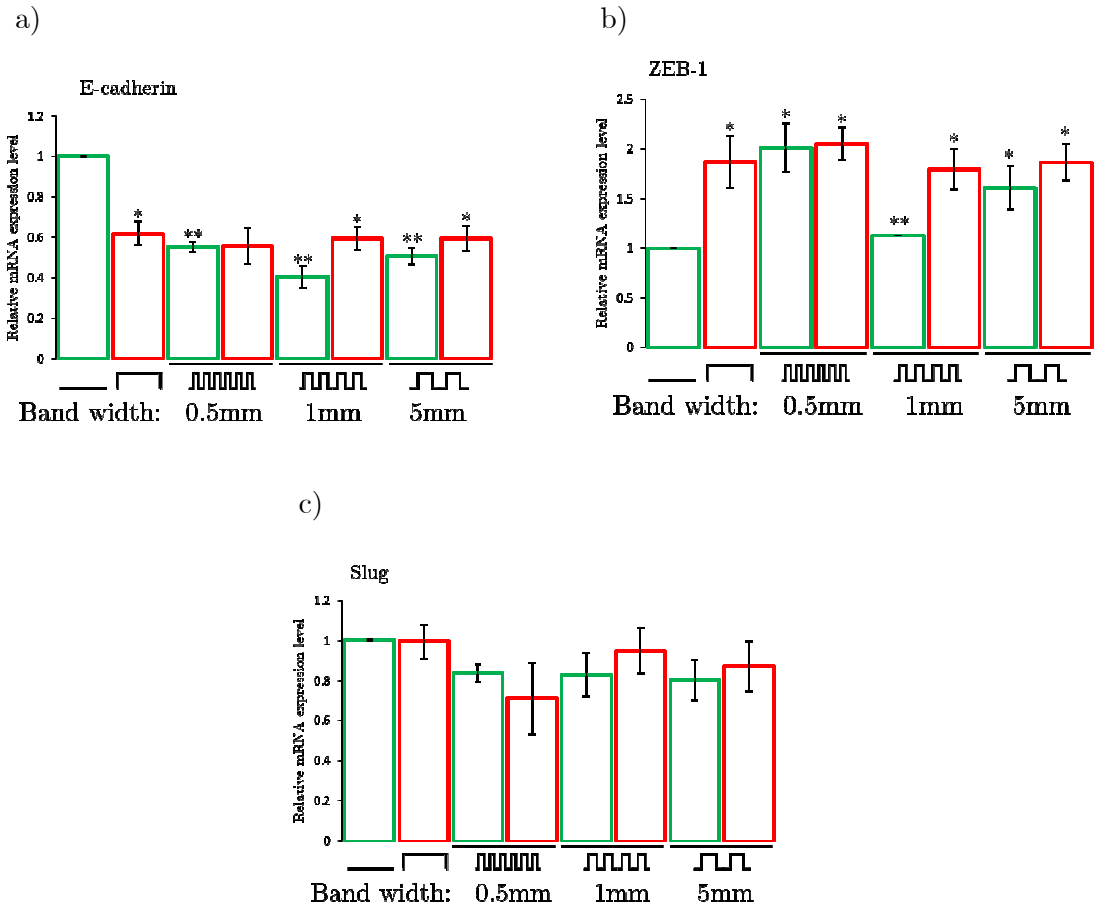


Figure 8.3 mRNA expression levels of EMT markers following band patterned distribution of EMT

MDCK-SnailERT2 H2B Dendra2 cells were either incubated in DMSO (negative control, or green disk) or in 10μm caged cyclofen for 1hr than washed (positive control and experimental conditions). The monolayer was illuminated for 7min at I_0=28mW/cm^2 with a 365nm UV light source and cells were washed again. Cells were incubated for 24hrs and then FACS sorted. mRNA from the green and red cell population was purified and qPCR analysis were performed on E-cadherin, ZEB-1 and Slug. Graphics below the histogram represent the pattern of illumination, numeric informs on the dimension of the bands. Color code: green represents the green cells,

or those that have not been illuminated. Red represents the red cells, or those that have been illuminated. Error bars represents SEM, n=5, p≤0.05.

Interestingly, E-cadherin and ZEB-1 levels are significantly affected while Slug does not seem to be responding. More specifically, the homogeneously induced condition (positive control) shows strongly reduced levels of E-cadherin and increased levels of ZEB-1 compared to the negative control. Levels of E-cadherin are strongly repressed in almost all conditions. In fact, the red cells in the 0.5mm band condition are not marked as significantly affected, but it is worth mentioning that the p-value associated to this condition is 0.06, very close to the cut-off value of 0.05. Similarly, all experimental conditions show significant increase in the mRNA expression levels of ZEB-1. Taken together, it appears as if, in all three experimental conditions, the green cells are not capable of remaining in the low ZEB-1 state of epithelia due to their coexistence with red cells. In other words, EMT induction in the red cells generates propagative influence to the neighboring green cells, forcing them to become more mesenchymal-like. At this stage, it is unclear whether the green cells are losing their epithelial phenotype as part of an active or a passive mechanism: are the green cells engaging into an EMT that is molecularly triggered by the red cells, or are the green cells losing their epithelial-like phenotype because their neighbors are disengaging from intercellular junctions?

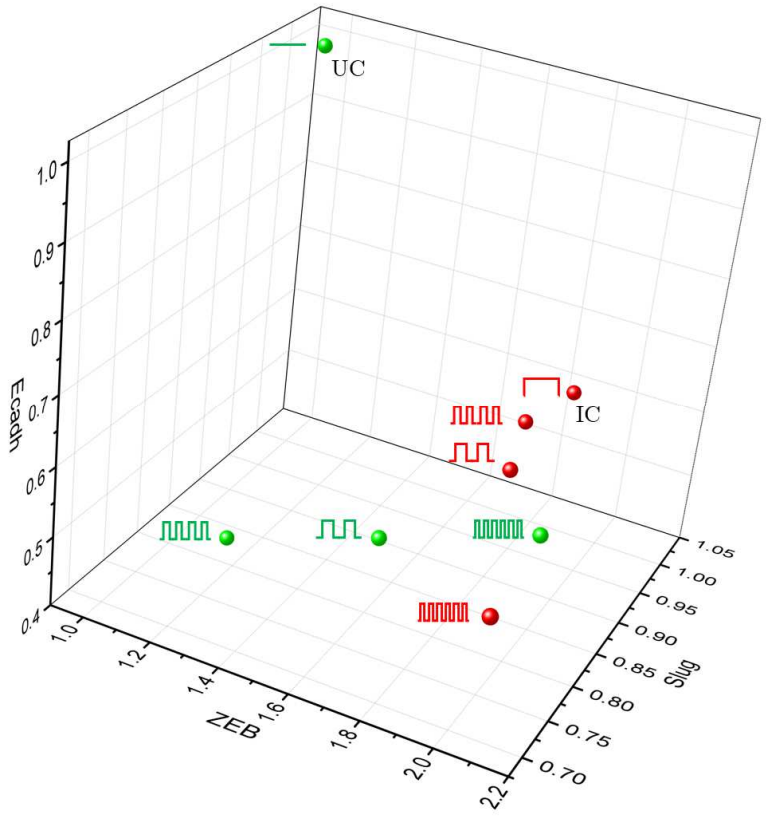


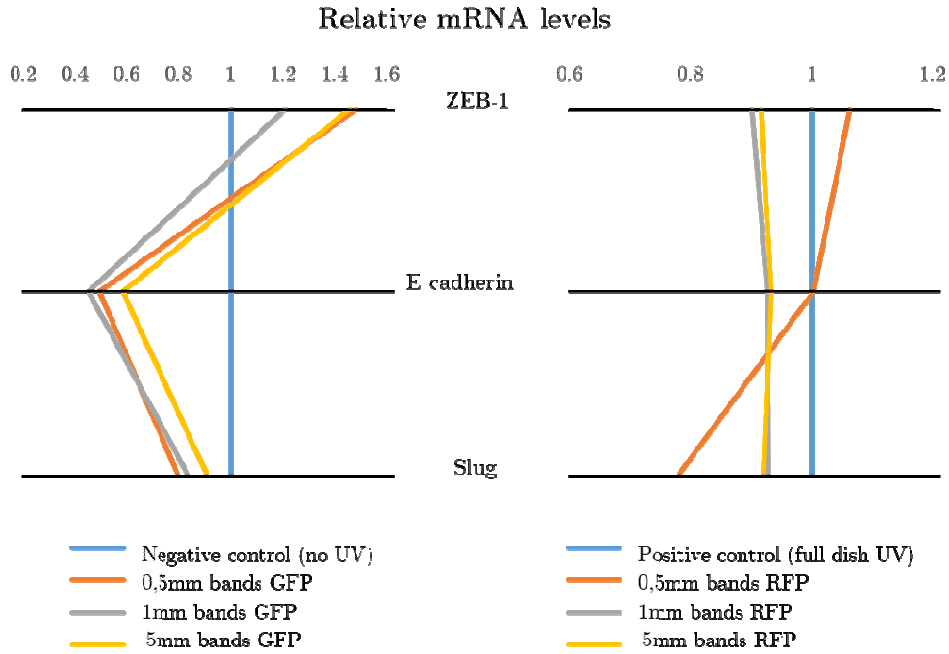
Figure 8.4 3D representation of mRNA expression levels of EMT distributed in symmetric bands

E-cadherin, ZEB-1 and Slug mRNA expression levels were represented on a single 3D graph. This representation brings out the existence of emerging properties. (UC: homogeneously un-induced control, IC: homogeneously induced control)

The 3D representation of the data allows to further appreciate the effect of the red cells on the green population. In fact, the geometric representation in the transcriptional state space show that there is an unanticipated attraction of all the experimental conditions towards the transcriptional state associated with the positive control. In this situation, where green and red populations are equal in size, it appears quite strikingly that the two populations are involved in a non-symmetrical competition where red cells have stronger influence on green cells.

The interdependence between the two populations can be visualized in Figure 8.5 where each experimental conditions is ratioed to its corresponding homogeneous control for the three genes. In other words, the green cells are compared to the negative control, and the red cells to the positive control. The blue vertical line refers to what would be expected if activated and inactivated cells were independent.

a)



b)

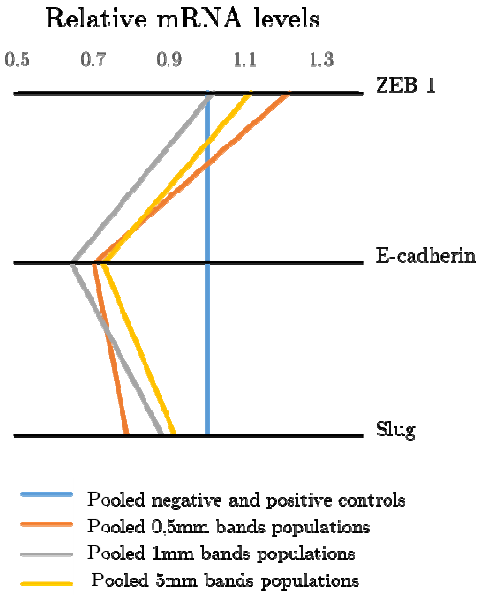


Figure 8.5 Representation of cell population interdependence in equal sized bands

mRNA data from the equal sized band experiment was plotted such that the green cells were compared with the negative control and the red cells were compared with the positive control. This was done for all three genes studied (a). Data shows that the green cells are very different from their control. Pooling the green and red populations for each conditions reflects the existence of emerging properties which explain why the heterogeneous

experimental conditions do not align with their control (blue vertical line) (b).

If the induced and non-induced populations were independent one from the other, the former should be comparable to the positive control whereas the latter should not be different from the negative control. It is quite obvious from Figure 8.5, a) that this does not hold true. In fact, green cells which are expected to have an epithelial transcriptomic profile are strongly divergent from the negative control, suggesting that these cells are less epithelial than expected and that they are strongly under the influence of neighboring red cells. On the other hand, red cells, which should be mesenchymal are, to a certain extent, reasonably comparable to the positive control. Similarly, if the two populations were fully independent, then the average of green and red cells for a given condition should be comparable to the average of the negative and positive control. The fact that this is not true (Figure 8.5, b) proves that the two populations are not independent and reflects the existence of emerging properties.

8.1.3 Non-equal sized band patterns

Given the previous result, the monolayer was then challenged to a similar band motif of EMT distribution. However, in this case, the 1:1 ratio was replaced by a 2:1 ratio of green over red cell size population. In fact, the previous result show that in equal sized populations, the red cells were more potent at imposing their phenotype to their green counterpart. By generating patterns where the green cells are now twice as numerous as the red cells, this experiment will address the question of parameters that may be important in understanding emerging properties. The patterns generated consisted in bands of 0.5mm illuminated/1mm non illuminated and 1mm illuminated/2mm non-illuminated. The band motifs are schematically represented below the histograms. The negative control is represented as a straight line whereas the positive control is an upside down U shaped symbol. The color of the histogram band reflects whether cells have been illuminated (red) or not (green). Due to the size of the well, the 5mm band condition from the previous experiment could not apply here as the pattern would be larger than the actual well.

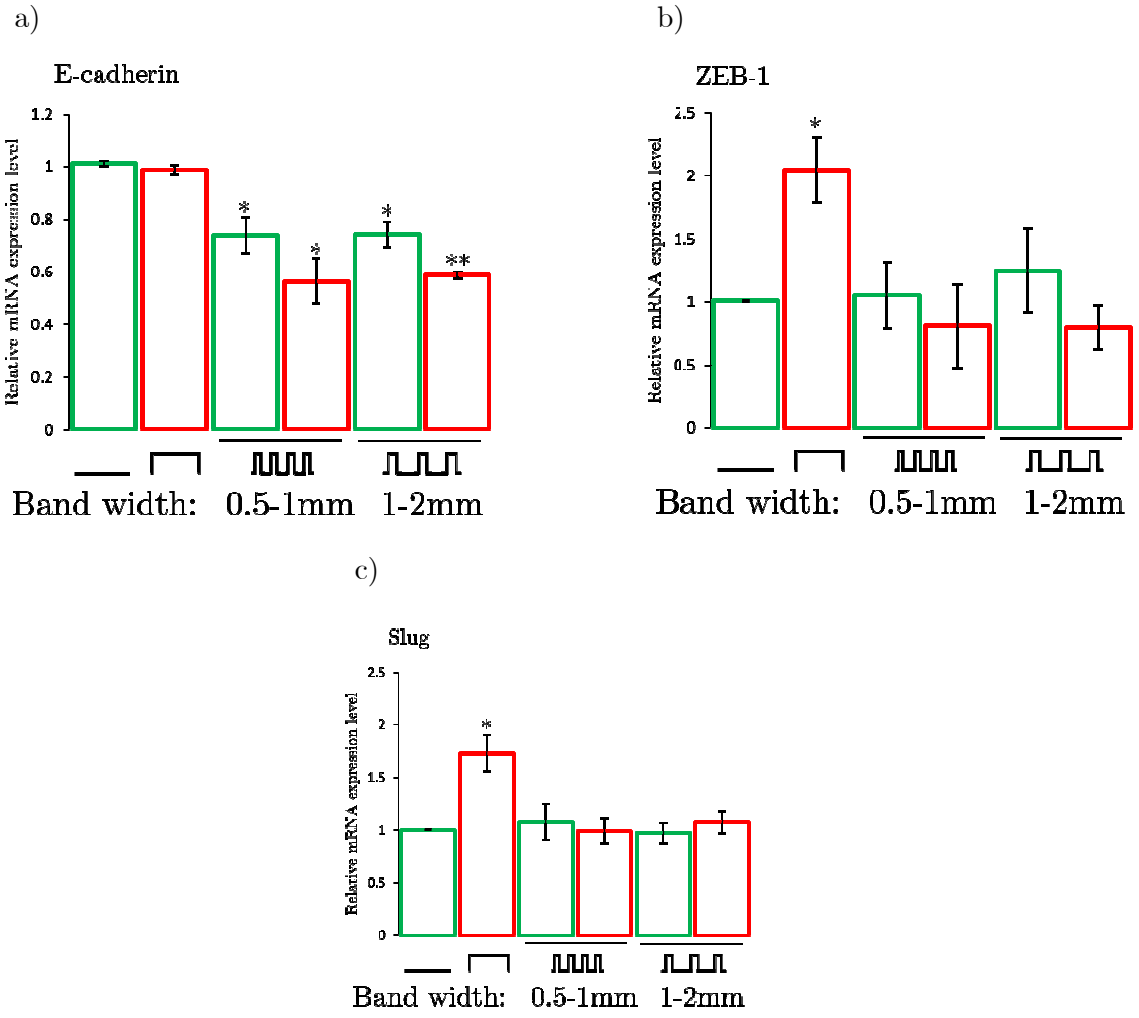


Figure 8.6 mRNA expression levels of EMT markers following non-equal band patterned distribution of EMT

MDCK-SnailERT2 H2B Dendra2 cells were either incubated in DMSO (negative control, or green disk) or in 10μm caged cyclofen for 1hr than washed (positive control and experimental conditions). The monolayer was illuminated for 7min at I_0=28mW/cm^2 with a 365nm UV light source and cells were washed again. Cells were incubated for 24hrs and then FACS sorted. mRNA from the green and red cell population was purified and qPCR analysis were performed on E-cadherin, ZEB-1 and Slug. Graphics below the histogram represent the pattern of the illumination bands, numeric inform on the size of the bands. Color code: green represents the green cells, or those that have not been illuminated. Red represents the red cells, or those that have been illuminated. Error bars represents SEM, n=3, $p \leq 0.05$.

Cells were induced as previously, and the results obtained when the two co-existing populations are no longer equal in proportion is very different from when the competition is symmetric in number. In fact, when the green population is twice as large

as the red one, none of the experimental conditions show significant increase in either ZEB-1 or Slug expression levels although the positive control is significantly affected (Figure 8.6). The levels of E-cadherin are ambiguous to interpret as the experimental conditions appear to be responding but the positive control is not affected in this case. In addition, it is unclear why in Figure 8.3 both E-cadherin and ZEB-1 levels are significantly affected while in Figure 8.6 E-cadherin levels are not affected, but Slug expression is. It is important to insist on the fact that in both experiments, ZEB-1 levels are responding. This is critical due to the central role of ZEB-1 in the induction of EMT. In other words, although the behavior of E-cadherin and Slug is not clear cut, the fact that ZEB-1 is constitutively affected in all the positive controls reaffirms that EMT is taking place and that interpretation can be done on the various experiments. The data obtained in non-symmetrical distribution can, as previously, be presented on a 3D scatter transcriptional state space plot to appreciate the response as a whole (Figure 8.7).

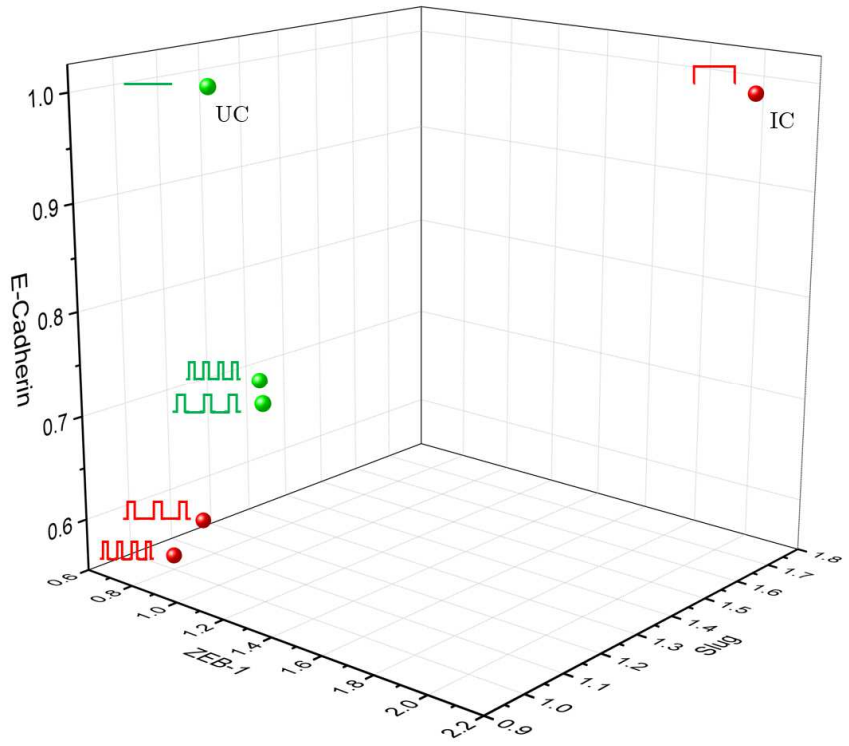


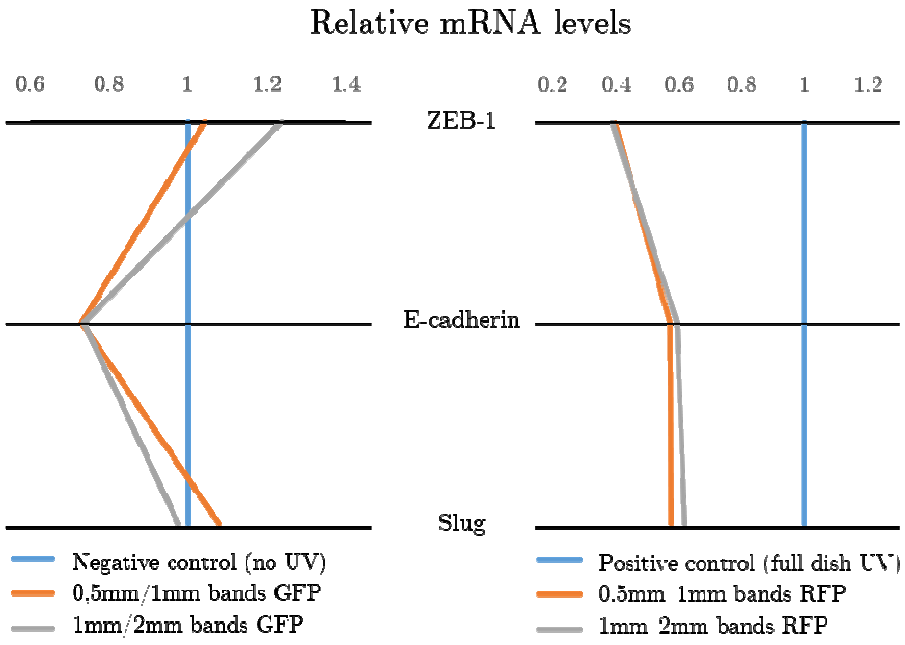
Figure 8.7 3D representation of mRNA expression levels of EMT distributed in asymmetric bands

E-cadherin, ZEB-1 and Slug mRNA expression levels were represented on a single 3D graph. This representation brings out the existence of emerging transcriptional responses. (UC: homogeneously un-induced control, IC: homogeneously induced control)

When comparing the negative and positive control, there is a net shift in the transcriptional state space between the two as a result of the significant increase in the expression of ZEB-1 and Slug, as mentioned previously. E-cadherin levels remain relatively high when all the cells are induced. In general terms, in the transcription state space, the different experimental conditions are all spatially confined around the negative control and red cells are very distant from their expected transcriptional localization near the positive control. This 3D representation brings out the striking effect of co-existing populations of unequal size. In fact, when green cells are twice as numerous, the situation is reversed compared to Figure 8.4 as they now seem to have a stronger influence on the fate of the red cells, forcing them to adopt a more epithelial-like, or less mesenchymal-like transcriptional profile.

Similarly to equal sized bands, the degree of interdependence between the two populations can be assessed by comparing each experimental condition to its corresponding control. When the GFP population is twice as big as the RFP population, the green cells have a transcriptomic profile that is comparable to the negative control. On the other hand, the red cells are now strongly divergent from the positive control (Figure 8.8, a) suggesting that in this configuration, green cells overweight the competition. Again, pooled populations of green and red cells differ greatly from the pooled negative and positive control revealing that two co-existing populations cannot be studied as the sum of these two individual populations.

a)



b)

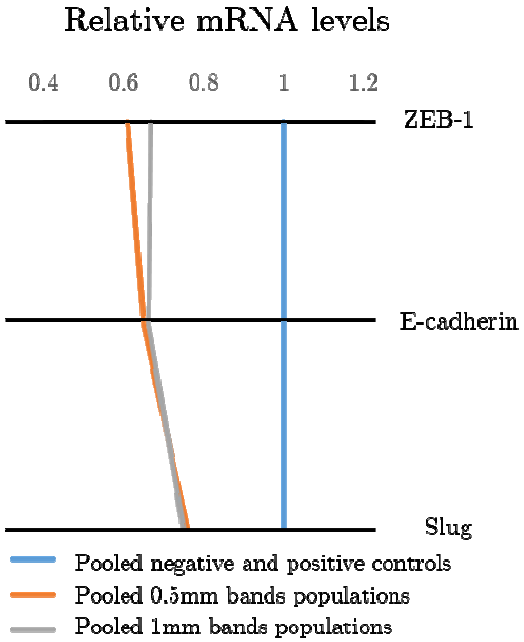


Figure 8.8 Representation of cell population independence in non-equal sized bands

mRNA data from the non-equal sized band experiment was plotted such that the green cells were compared with the negative control and the red cells were compared with the positive control. This was done for all three genes studied (a). Data shows that the green cells are very different from their control. Pooling the green and red populations for each conditions reflects the existence of emerging properties which explain why the experimental conditions do not align with their control (b).

Given these result, it can be stated with confidence that two co-existing populations do not behave independently one from the other, but in fact influence their average fate, as measured by the transcription response. Whether this influence is the result of molecular hijacking of the cell or of direct mechanical effect is yet to be understood. Suffice to say that the outcome of heterogeneous interactions can be modulated by the size of the two opposing populations, affecting which population becomes dominant over the other. This data also shows that the epithelial phenotype is not under cell autonomous regulation but depends on whether neighboring cells promote, or not, the required organization for its existence.

8.1.4 Single cell induction of EMT

Given the interesting results obtained when heterogeneous EMT is induced, it became necessary to reduce the size of the patterns used in order to progressively "dissect out" the rationale of emergent properties and to further study the behavior of the single cell rather than the average behaviors of green cells and red cells. In fact, varying parameter on populations of much smaller size should render the understanding of epithelial stability more straightforward. Using *Selective Illumination* described earlier, EMT was induced in a monolayer according to strict geometrical rules.

Cells were incubated in caged cyclofen and illuminated using the 405nm laser beam and the FRAP head described previously (Section 7.4.3). The laser power was set such that the fluence was comparable to the value found in Section 7.2.2 on a region of interest (ROI) of known and reproducible area.

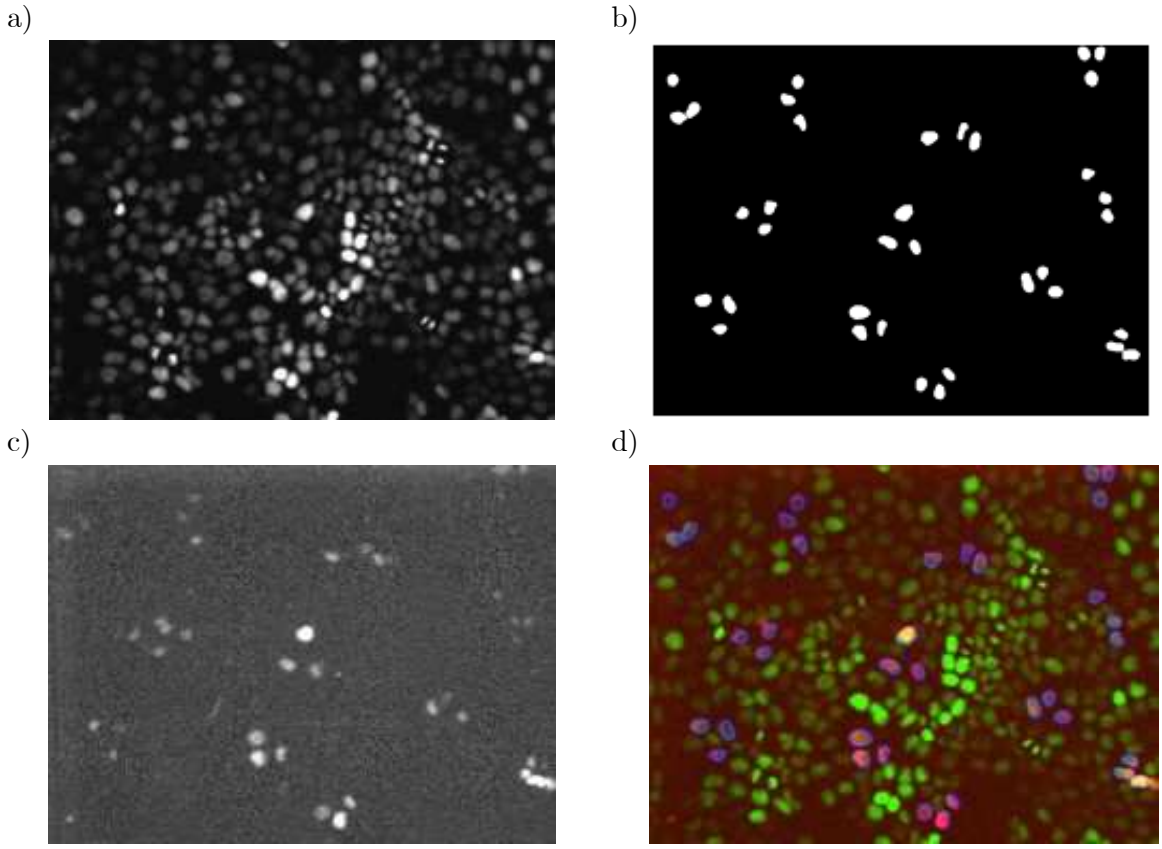


Figure 8.9 Local induction of EMT using *Selective Illumination*

MDCK-SnailERT2 H2B Dendra2 cells were bathed in 10µm caged cyclofen for 1h then washed. Cells were illuminated using the 405nm laser beam and the FRAP head. Positions to be illuminated were determined using *Selective Illumination*. The parameters of the program were set such that the cluster size = 3 cells and the distance between each illuminated clusters = 1 cell. The image of the nucleus (a) was used to generate the list of positions to illuminate (b). FRAP illumination on these positions leads to photoconversion of the Dendra2 protein as well as cyclofen uncaging (c). The overlay image shows the initial nucleus (green), as well as the contours of the regions to illuminate (blue empty circles) and the actual positions where the illumination was done (red) (d).

This experiment was done at the very last stage of this thesis work. As a result, FACS sorting of the two populations was done, but transcriptional analysis remains to be completed, and raises the issue of the need for a direct transcription reporter that can be read *in situ* in the cell.

Although no data has been generated yet, it is important to mention that this approach is extremely powerful as it allows to generate illumination that is interactive and "monolayer specific" in the sense that it takes into account direct neighbor patterns. Also, because these experiments are done in a dish, and the illumination does not cover the whole dish but rather allows to cover several field of views only, data from the green

cells will not be exploitable. Indeed, this experimental limit will generate strong heterogeneity within the green cell populations: green cells in direct contact with red cells will be jointly analyzed with green cells outside of considered fields of views. However, this technical limitation can be overcome by growing a monolayer with controlled size. This has been tested, and can be achieved through photolithography patterning methods [147]. Together, this now functional and precise method can be used to vary cluster size and minimal distance to study how, for each initial situation, the monolayer responds as a whole.

Chapter 4

IV. Discussion

9 The experimental model

9.1 An opened choice

The present work aimed at studying the central problem of cellular interdependence, from a viewpoint and with a cellular model that both need to be discussed. The goal is to ultimately be able to assess the extent of cellular autonomy, i.e. the relative importance of internal vs. external determinants in a manner that these two sources of information be quantifiable. To make the question more tractable, the consideration was made that the extent of autonomy vs. interdependence should reflect on the effect caused on the cell behavior by non-congruent variations of the internal and external determinants, i.e. when these variations are expected to lead to opposite consequences. In the context of the epithelium, this idea is embodied in the situation where one cell receives an EMT trigger that modifies its internal state, while the neighboring cells do not; because the non-activated neighbors possibly contribute to a normalizing effect, the perturbed cell is under two competing information sources: undergo EMT because of the EMT trigger, or remain epithelial because of the possibly normalizing neighbors? Such a scenario where non-congruency is generated to measure cellular interdependence forms the basic framework of this thesis. Two main questions thus drive this work: could an individually-delivered cause lead to different effects depending on the geometry underlying the distribution of that cause among the cells of a community? Could logical rules of interdependence be inferred from how the cause-effect relation depends on the geometry of the cause?

The way this problem is framed is related but somehow different from a large number of studies dedicated to understanding how cells mutually control each other within a community. Indeed, in most studies of cellular communication and

interdependence, the focus is to understand the mechanisms leading to well defined patterns of differentiation mostly during development. There is extensive data on the reproducibility and robustness of pattern formation in *C. elegans* vulval development [148, 149], as well as in pattern formation of the imaginal disks in *Drosophila melanogaster* development [150]. In both cases, the emerging patterns are well described and studied from the point of view of the mutual logics of the decision making processes, and how position information is robustly produced and handled. In contrast, the model studied in this work is composed of an apparently homogeneous and phenotypically steady state population of fully differentiated epithelial cells, and not much is known so far to happen in terms of differentiation or phenotype decision making. Based on common thoughts, there is no reason to question why an epithelial cell remains epithelial nor why stable junctions remain stable. Incidentally, a previous thesis work from Simon de Beco showed that in fact junction stability is a problem, and does require a mutual active stabilization machinery [110]. Here, the mere consideration that the epithelial phenotype cannot be autonomously reached, leads to the hypothesis that epithelial stability is intrinsically a collective process. The latter relies on interdependence rules that are defined by geometry, logics, mechanics and biochemistry. It is important to mention here that a number of studies have considered the collective mechanical behavior of epithelia during cell motility and wound healing [60, 151, 152], but the issue of phenotype stability as an emergent process has not been studied so far. In the situation of developmental EMT or epithelial dysplasia which both challenge the stability of the epithelial phenotype, the main line of past investigations has been to look for molecular mechanisms, but the effect of heterogeneity observed in clinical settings has not been studied *per se*.

9.2 An optical approach with a will defined unitary cause

Because epithelial stability must be addressed from the starting point of an epithelium that is fully differentiated and as homogeneous as possible, cellular models based on mixing different cell types such as mixing EMT-inducible and non-inducible cells had to be ruled out. In this respect, the optogenetic tool used in this work specifically allows to carry out experiments in an initially homogeneous system.

To study the degree of cellular interdependence, the model relies on the deposition of a unitary cause in individual cells. This cause, the induction of EMT through photoactivation of cyclofen and activation of Snail1, is invariable and reproducible. In fact, by definition, all the cells of a given sample are exposed to 10µm caged cyclofen. Because the illumination dose has been calibrated and is fixed, each individual cell will

receive the same amount of photons as its counterparts, suggesting that the fraction of active cyclofen will be similar within illuminated cells. Slightly varying expression levels of the fusion protein SnailERT2 between cells is a potential source of variability that could affect the degree of response of a cell. In addition, due to the nature of the illumination masks, it is likely that cells at the border of the mask receive partial illumination. This is a potential source of heterogeneity which should be overcome in the next experiments with the *Selective Illumination* approach.

The experimental model is functional but it is yet important to question whether the chosen approach is the most favorable. Indeed, here, EMT is induced through activation of Snail1. How biologically relevant it is to isolate Snail1 as a unique inducer? The fact that transcriptional data show significant changes in ZEB-1 expression levels is convincing to state that this simplified induction is yet functional. However, it will be interesting to see how this induction could be further optimized. As mentioned in the results, there is some debate on the kinetics of nuclear translocation and activation of Snail1 which is further blurred in this experimental model because it is part of a fusion protein. A better understanding of the kinetics of the induction would possibly allow to force Snail1 to the nucleus in a more robust manner such that its activation could be stronger. In addition, the central role of ZEB and TWIST family members has been discussed earlier, and it would be interesting to assess whether activation of all three EMT-TFs families could induce a more drastic EMT that would be more straightforward to detect and to quantify.

In general terms, EMT is an extremely complex process, home to many unanswered questions and debates. The fact that this project is based on such a complex event has inspired here a very simplistic experimental model, where EMT was arbitrarily chosen to be induced in a manner than may not include the whole picture. This simplification is not an issue since the details that make up the molecular complexity of EMT are not the main focus of the project at this stage of advancement. The question of emerging transcription patterns and social underlying rules of epithelial stability constitute a great void in our current knowledge and remains a broad and opened quest. It is however important to keep in mind that, as these rules are revealed, the experimental model will need to be challenged and rethought to allow a detailed understanding of underlying molecular mechanisms.

9.3 How to best read an EMT response?

The choice of the ideal readouts was a main concern at the onset of the project: EMT is a multidimensional process known to involve a large number of well-defined

markers [42]. Because the scope is not to characterize the molecular mechanisms of EMT but rather use it as a tool to generate heterogeneity within an epithelium, it was not of critical importance to fully recapitulate the molecular signature of EMT. Rather, it was key to find a readout that would be sufficient to say with confidence that the SnailERT2 fusion protein has been induced and that the cascade of events leading to a phenotypic switch has been activated. Whether all the markers of EMT were responding in the conventional manner was not critical. For these reasons, the study of the expression of ZEB-1, Slug and E-cadherin which is certainly incomplete in the face of the many known markers of EMT, is sufficient, in this model, to appreciate with confidence that the phenotypic switch has indeed occurred. Coupled to morphological images and immunofluorescence, this ensemble of readouts are relevant and sufficient at this stage of the work.

A straightforward reporter of an event is a binary readout. In fact, binary systems in which elements exist only in two states are ideal when trying to simplify rule analysis. In particular, a binary readout would make it easier to study phenotype metastability and instability nucleation rules, which is an interesting question in the context of carcinogenesis. However, the EMT transcriptomic response is known to be scalar and, as previously mentioned, several reports in fact show that EMT is a multidimensional process [62, 63]. As a consequence, making binary decisions based on the transcriptional levels of genes could thus be too simplistic and highly questionable.

A possible binary readout that was considered in this project is the primary cilium of epithelial cells. This solitary organelle is present at the apical surface of epithelial cells and is thought to play a key role in sensing and transducing signals from the microenvironment [153]. The cilium is present only in quiescent cells, and has been shown to resorb itself when the epithelial phenotype is challenged through cell division, injury, or EMT [154, 155]. The binary aspect of this readout is highly relevant given that what would be assessed is the morphological presence or absence of the cilium rather than a graded transcriptional level. Immunofluorescence were performed on the MDCK cellular model, but unfortunately no primary cilia were detected (data not shown). This can be explained by the fact that for cultured cells to be strongly epithelial, it is important that they be grown on permeable substrates with adequate rigidity. These allow for the presence of medium on either side of the cells, necessary for full apico-basal polarity. In the current model, the MDCK cells are grown on plastic which allow a certain degree of epithelialization yet not sufficient for the primary cilia to appear.

9.4 Key properties of cell-cell interactions

Localization vs. non-localization

Cell-cell interactions can primarily be local or non-local. Local interactions include cell-cell adhesion, contraction mediated mechanical interactions, paracrine effects, and message secretion at time-scales where the message has no time to equalize over the entire cell community. If non-local interactions dominate over local effects, a simple consequence is that the outcome of a heterogeneous EMT induction should strongly depend on the relative population of induced vs. non-induced cells. Another consequence of non-local interactions would be that the response to heterogeneous perturbations is independent on the spatial scale of the heterogeneity. In line with this reasoning, the signature of dominantly local interactions should be that the response would strongly depend on the scale of the induction heterogeneity. Ideally, a position dependent analysis of individual cell responses should be informative on local vs. non-local effects.

Temporality of cellular interactions

Virtually all biological responses that are autonomously controlled by individual cells take time for obvious reasons. In addition, the mediation of chemical cell-cell interactions evidently also requires time because of message diffusion. As a consequence, the time frame of possible congruent or antagonistic messages should play an essential role. In the current cellular system, it is most certain that there is delayed transcription responses as well as delayed message transfer between cells. Because there is no knowledge about these contributions, the temporal aspects of the problem were kept out of the scope of this work. The minimal stand point was to operate synchronously over the cell population and fix the temporal parameters based on simple practical considerations: the time frame should be sufficient to allow a detectable EMT event to occur while avoiding long times at which the Dendra2 readout vanishes. Although this approach is sufficient for the time being, future work could better address this issue of temporality.

9.5 Photoactivation patterns and interdistances

The patterns used to induce heterogeneous EMT were generated according to various geometries. The goal was to vary the relative proportion of the populations as well as the scale of the heterogeneity (cluster size).

If the response of a given individual cell depends on the ensemble of all the messages integrated over time from all other cells, it will depend on the distribution of interdistances between this given cell of interest and all the other cells. In discrete

systems such as the extreme situation of the vulva in *C.elegans,* each cell is unique, with a unique set of interdistances. In contrast, in the geometry of an infinite chess board pattern of induced and non-induced cells, each cell is subjected to the exact same set of interdistances as any other cell of the same population. In intermediate situations such as those generated in this work, the symmetry of the activation pattern leads to distributions of interdistances that vary among cells (Figure 9.1).

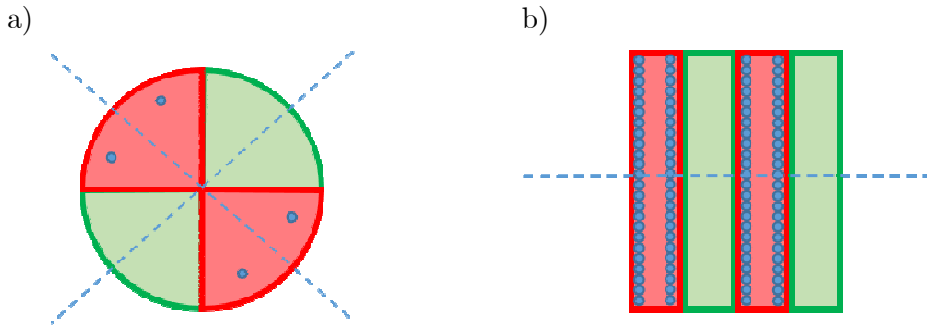


Figure 9.1 Interdistances in patterns

For each generated pattern of induction, a distribution of interdistances defines the neighboring landscape of a given cell. These distributions can be more or less heterogeneous depending on the geometry of the pattern. In radial distribution (a), there is a high degree of heterogeneity leading to a very small number of cells being exposed to the same distribution of interdistances. Band patterns (b) introduce periodicity in one dimension and thus reduce the heterogeneity of interdistances distributions, leading to a greater number of equivalent cells. (Red and green represent two different populations, and blue circles represent individual cells that have the same distribution of interdistances.)

Reducing the heterogeneity of the distributions of interdistances among green and red cells is critical given that the experimental approach used in this project is based on averaging the transcriptional signature over green and red populations.

9.6 Main observations

9.6.1 EMT in the transcriptional state space

Figures Figure 8.2, Figure 8.4 and Figure 8.7 propose a 3D representation of the transcriptional state in the different experimental conditions. Here, each point represents the average of cells within the green or red population. Assuming that the transcriptional state of individual cells is homogeneous, then representing individual cells rather than the average of cells should be equivalent: all red cells (and respectively green cells) should be located at the same average red (and respectively green) point. In other words, all un-

induced cells should be located on the average of homogeneously un-induced point (UC), and all induced cells should be located on the average of homogeneously induced point (IC). Secondly, assuming that the single cell response is binary, i.e. either switching to a fully mesenchymal state (IC point) or remaining epithelial (UC point), the average of green cells and the average of red cells, should necessarily be located on the [UC, IC] segment. This is obviously not the case (Figure 9.2), as the average of green and red cells are located significantly away from this reference segment. Rather, the data strongly suggests that heterogeneous EMT induction leads to emergent transcriptional states, and thus to a non-linear response in individual cells.

It is important to recall that these emergent transcriptional states are measured as population averages. In these conditions, it is difficult to known what could be the transcriptional state of individual cells inside these populations. As mentioned in Section 9.5, heterogeneous induction pattern lead to heterogeneous neighboring landscapes where each cell, or group of cells, is subjected to a given distribution of interdistances. Consequently, red and green points reflecting heterogeneous induction on the 3D graph probably represent the average of broadly distributed individual states.

The detail of individual responses is obviously very interesting and needs to be understood to reveal emergent individual responses. This is currently a limit of this work and will be studied in the near future.

9.6.2 The importance of proportions

The 3D representation of the transcriptional state using ZEB-1, Slug and E-Cadherin allows to reveal and measure the distance between any two points. These distances are represented in Figure 9.2 and can be used to reveal the degree of influence of two co-existing populations.

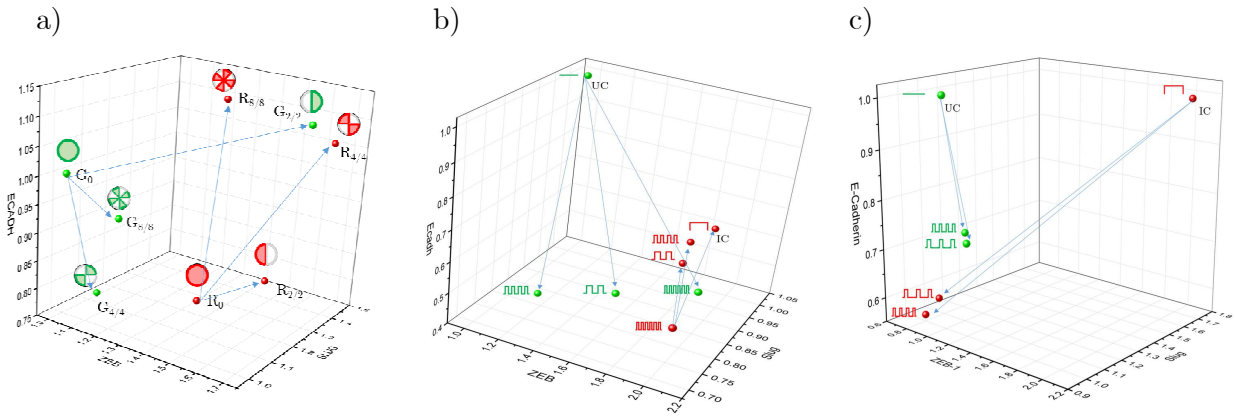


Figure 9.2 Degree of influence of populations

Figure adapted from (Figure 8.2, Figure 8.4 and Figure 8.7) showing the distances between a given population part of heterogeneous EMT induction and its relative homogeneous control in (a) radial induction, (b) symmetric induction and (c) non-symmetric induction. The size of the distances reflect the degree of mutual influence of two co-existing populations.

In radially distributed induction (Figure 9.2, a), how much non-induced cells are influenced away from their homogeneously induced relatives can be assessed by the length of the state-space distances $[G_0, G_{2/2}]$, $[G_0, G_{4/4}]$, and $[G_0, G_{8/8}]$. Large influences are reported on Figure 9.2 (a), but the interpretation remains difficult for reasons that have previously been mentioned (Section 8.1.1). In the situation where induced/non-induced cells are equally proportioned (Figure 9.2, b), the distances between states show that the influence of induced cells on non-induced cells is much larger than the reciprocal: distances between heterogeneously induced green cells and their homogeneous non-induced control are much larger than distances between heterogeneously induced red cells and their homogeneous induced control. When the population balance is biased towards green cells in a 2:1 ratio (Figure 9.2, c), the strength of this influence is reverse: distances between red cells from heterogeneous induction and their relative control are much greater than those for the green cells.

These results provide experimental proof of biological non-linearity: co-existence of EMT induced and non-induced cells lead to a mismatch between the expected phenotype of a population and its actual transcriptional state. This non linearity can only be explained as the consequence of mutual influence in heterogeneous situations, and highlight the existence of interactions which individually alter the cause-effect relationship.

9.6.3 Local vs. non local effects

Following the earlier discussion on locality (Section 9.4), it is very unlikely that the influence of induced cells on non-induced cells, and *vice versa*, would be entirely non-local. Nevertheless, the present data does not give any clue to totally rule out this possibility: the positional information of each individual cells is totally lost in the qPCR and sorting analyses. It is however interesting, and necessary, to discuss about possible local interactions for the mere reason that EMT *cannot* be a single cell process. By analogy with the nucleation barrier in 1st order phase transitions, it is possible that a small enough cluster of induced cells is frustrated and remains fully epithelial, but that large enough clusters undergo the transition and influence neighbor cells.

In the experiments, because the green and red domains are large, only a small proportion of cells lay at the induced/non-induced boundaries. As a result, the local

contribution of cells at these junctions is probably diluted out in the average signal. To evidence EMT nucleation barriers, more studies are needed with local reporters.

Finally, discussing local vs. non local effects is not only limited by the absence of local observables in these experiments, but also by the fact that the observation is not dynamic and not performed at short enough times. Indeed, the photoactivation triggers a complex cascade of events, leading to a hypothetical scenario. First, the induced cell will lead to the activation of a cascade of snail-induced responses which include early internal and autonomous events. The data shows that cells secondarily deliver the signals to non-induced cells. Immediate neighbors are likely to be influenced first, primarily because of their junctions shared with induced cells, but also possibly because of paracrine communications. Beyond the nearest neighbors, second, third and more distant neighbors will then be influenced. The role of time in spreading the influence at longer distances strongly depends on the spreading mechanism (mechanics, diffusion, etc...). Regardless of that mechanism and its kinetics, the experimental data shows that cells respond similarly despite the geometry of the pattern (different band width) (Figure 9.2 b and c). This strongly suggests that the influence has propagated fast enough to yield a length-scale independent effect within the observation time. Larger structures and/or shorter observation times are needed to explore the propagation of the influence.

10 Perspectives

10.1 Following experiments

10.1.1 Conditioned medium

The literature provides growing evidence for the role of soluble molecules secreted by cells that act as potent effectors in cellular competition and collective behaviors [92, 156]. These effectors can be growth factors, hormones, exosomes or microRNAs (miRs)[92, 157]. To better characterize the non-local effect of heterogeneous EMT induction, conditioned media experiments were performed. It is known that cells induced to undergo EMT will, beyond cytoskeletal reorganization, increase the relative expression of certain miRs [158, 159]. With this in mind, it is interesting to speculate that soluble factors secreted by an induced cell could potentially influence a non-induced cell. This type of influence *via* the medium can easily be assessed. Conditioned medium experiments consists in priming cells to undergo EMT in a homogeneous manner, and collecting the medium of these cells to place it onto non induced cells. If in fact the EMT response is influenced by factors contained in the medium, the transcriptomic profile of

these non-induced cells should be affected and appear more mesenchymal. The reverse can also be performed where induced cells could hypothetically be reversed when exposed to conditioned medium collected from non-induced cell cultures. These experiments were carried out to test this in both homogeneous (all the cells) and heterogeneous (band patterns) conditions. At the time of submission of this manuscript, the experiment was performed three times, and still needs to be repeated twice in order to conclude significantly. Hopefully by the time of the defense this data will have been collected and analyzed.

10.1.2 Local readouts

The experimental model used has proven effective at revealing the existence of emerging properties. The limit of this model is that no individual cellular information can be recovered due to the loss of positional information. To overcome this, it would be feasible to improve the current model in order to generate local and dynamic readouts of the EMT response. For this, E-Cadherin, Slug and ZEB-1 could still be used, but rather than measuring their relative transcriptional levels by qPCR, a live reporter of their activity could be established. In fact, placing a fluorescent reporter downstream of the promoter of each gene would allow to quantify the level of expression from the measured level of fluorescence in an unperturbed monolayer, in a dynamic and spatially resolved manner. In addition, such an approach would overcome the issue of averaging intra population heterogeneity. Finally, the understanding of the geometry, logic and nucleation barriers would be more accessible. Coupled to the *Selective Illumination* tool, this live and local readout would allow to generate single cell information.

10.1.3 From to 2D to 3D biology

In the scope of pursuing the project further, it would be extremely relevant to study how the biology presented in this manuscript can be transposed into a more physiological 3D cellular model. Ground-breaking work from Mina Bissell in breast and ovarian cancer have revealed considerable evidences that cell-cell and cell-extracellular matrix (ECM) interactions play a crucial role in shaping the nature of the microenvironment [2, 9, 160, 161]. Furthermore, during the first "Goodbye Flat Biology" meeting, held in Berlin in 2014, the importance of using 3D cellular models for the understanding of biological event was strongly highlighted. During this meeting, many examples revealed that cellular processes taking place in a 2D configuration do not always mimic the actual 3D reality which characterizes the living [162, 163]. Luckily, MDCK cells, when grown in specific extracellular matrix known as Matrigel, have the capacity to form hollow 3D structures called cysts [164]. These multicellular structures have a well-established apical-basal polarity organized as a monolayer around a hollow

lumen. 3D MDCK structures recapitulate many features of *in vivo* epithelial tissues, and thus constitute an ideal model for the study of intercellular junctions under physiological conditions in a Petri dish [165]. Preliminary experiments were done to assess the feasibility of heterogeneous EMT induction experiments in 3D, taking into account technical issues involved with growing cells in Matrigel: incubation of the cyclophen, washing of the cysts, imaging etc... Cysts were grown until the central lumen was visible, and immunofluorescence was performed in control or TGF-β induced cysts. In the latter, the 3D structure was drastically affected causing loss of the organization of the monolayer and disappearance of the central lumen (Figure 10.1). Strong reorganization of the stained markers show pronounced loss in the cellular polarity. The true nature of the observed effect was not further analyzed as adequate experimental conditions were yet to be properly defined, including 2-photon photoactivation.

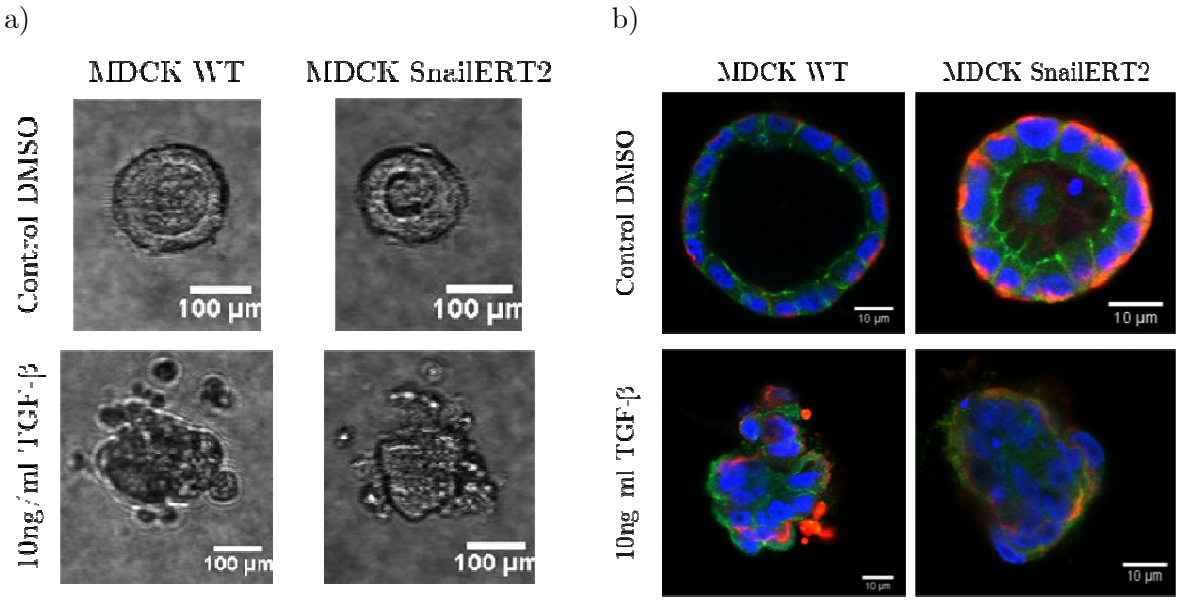


Figure 10.1 Morphological and immunofluorescence of MDCK cysts

Transmission images of MDCK WT and SnailERT2 cysts (a). Cysts were grown for 7 days in extracellular matrix and were incubated either in DMSO or 10ng/ml TGF-β for 48hrs. Drastic loss of the 3D conformation is observed. Immunofluorescence images (b) of these same conditions show similar loss of cysts organization. Staining mark E-cadherin (green) and Vimentin (red). The nuclei were counterstained with DAPI (blue).

The induction using cyclophen/caged cyclophen was also tested, but due to the high concentrations of proteins found in the Matrigel, the cyclophen molecules which non-specifically bind to proteins could not reach the cells before being sequestered by the ECM. A summer intern worked on this issue for three months, and optimized the

protocols for cysts handling, washing and induction without disrupting the 3D structure. The proper conditions have been established and this 3D cell model could now potentially be tested for heterogeneous induction of EMT.

10.1.4 Molecular mechanisms

The above improvements of the current experimental model would then allow to address the question of underlying molecular mechanisms that are driving the rules of epithelial phenotype stability. Similarly to the local and dynamic reporters described above, an even more complete set of readouts could be generated to not only to reflect transcriptional expression levels of certain genes, but also to fluorescently stain proteins of interest (ZO-1, Snail1, Twist, α-SMA and others known to play a role in EMT). These transcription networks entail challenging molecular biology steps but would prove to be extremely relevant to obtain dynamic and spatial data.

With these powerful tools, it would then be possible to look for candidate molecules that mediate the information from one cell to the other and targets could be silenced using siRNAs. Possible candidates could be miRNAs, homeotic genes, proteins involved at the cell junctions etc....

10.1.5 How does this work relate to cancer?

Carcinogenesis implicitly originates from dispersed oncogenic causes (spontaneous mutations) revealing the intrinsic heterogeneous nature of the process. Surprisingly, oncogenes have never been tested for their oncogenic properties in well controlled heterogeneous patterns, except in seminal experiments by Rubin [83, 166]. With this in mind, the present work could, in the future, be oriented towards the local photoactivation of oncogenes. This could be done with a cellular model that would be somewhat improved in terms of its physiological relevance (cell line, 2D vs. 3D, mechanics etc....).

11 Final remarks

The main point that strikes out in this work is that the territory, or neighborhood, of a cause breaks the univocal cause-effect relationship. This is also true in carcinogenesis which is known to stall in situations where the histology suggests a mismatch between heterogeneously distributed causes (mutations, viral infections etc...) and their expected consequence (transformation). The outcome of some dysplasic precancerous lesions can be regression, suggesting that the tissue is a battlefield between potentially cancerous cells and possibly normalizing "normal" cells. The control mechanisms at play in frustrating the effective transformation and maintaining an oncogenic source aphenotypic is key, and understanding its molecular and biochemical players can be extremely insightful. Indeed, oncogenic anarchy arises when normalization fails. The rules underlying mutual phenotype control in the epithelium should help one day to understand the sociology of the battlefield and to better predict the outcome of precancerous lesions.

Dysplasia are unavoidable, but their progression to malignancy can and should be controllable. This doctoral work has focused on revealing very preliminary aspects of cellular interdependence. By simplifying a fundamental biological question, it was possible to reveal a certain logic, or social rules, which could govern how a cell integrates its internal determinants and the signal received from its neighbors. Understanding how the cause-effect relationship can be distorted by the influence of the surrounding cells is key in predicting possible outcomes. Placed in the context of cancer, this work leaves many unanswered questions but has risen a thin veil and a new vision on the comprehension of oncogenesis. This work has laid the foundation of a very ambitious project which will hopefully lead to insightful future endeavors.

12 Appendix

12.1 Selective Illumination

12.1.1 Cell segmentation

In the following, the steps of the method are described in detail:

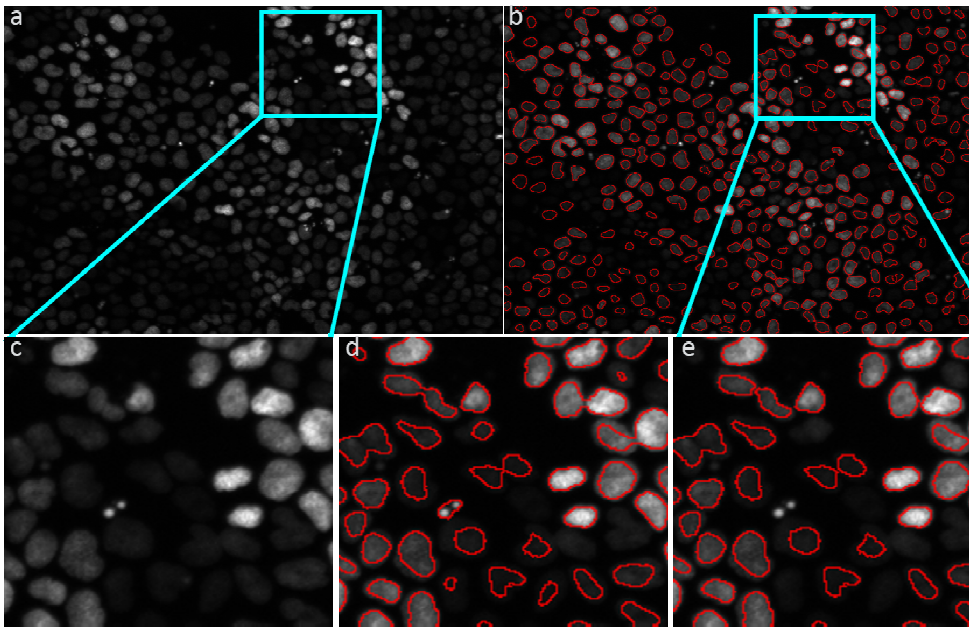


Figure 12.1 Segmentation steps

(a) Original image, (b) final segmentation result; cyan square is shown in subfigures (c-e), (c) original image (detail), (e) result of the segmentation prior to post-filtering and split (e) final result (detail).

(1)**Segmentation of nuclei:** The segmentation of nuclei is a well-known problem in microscopy image analysis. Here, the main difficulty is that the method needs to work with relatively low signal to noise ratio. Nevertheless, it was possible to solve this issue with relatively traditional tools: (i) for pre-filtering, we first apply a median filter (radius 2 pixels) to remove noise without impacting too much the object contours. We then applied a bilateral filter that smoothes the image, without blurring the object contours (sigma in space is chosen 2 pixels, sigma in grey levels is 0.15). Finally, morphological closing and opening by reconstruction are used to remove spurious dark and bright elements, respectively. The result of these pre-filtering steps is the prefiltered image p. (ii) Two binary images are

obtained by thresholding a background subtracted version of the image and the morphological h-maxima transform. For background subtraction, a local average in a fairly large window (here chosen to be 20 x 20; implemented with integral images to reduce computation time) is subtracted from the prefiltered image and the residue is thresholded at a fix and global threshold, as local background variations have been removed by the background subtraction (the chosen threshold is 4 in this case). The second binary image is calculated by the h-maxima transform which consists in first subtracting a constant (here h=10 was chosen) from the prefiltered image and then applying morphological reconstruction by dilation of the shifted version under the unshifted image. This removes white blobs from the image. Subtraction and thresholding (threshold = 4 was chosen) then results in a second binary image. The union of these two binary images is used for further processing. (iii) In the post-processing, holes are filled (morphological closing by reconstruction, size 11; the traditional hole filling algorithm was not applied as it also removes spaces surrounded by nuclei which therefore leads to problems for high cell densities). Then, close nuclei wrongly segmented as a unique object are split by the well-known Watershed algorithm, applied to the inverse distance map of the binary images. In order to avoid small contour irregularities to impact the segmentation result, morphological dynamics are used to preselect the minima from which flooding is initiated. Finally, objects with too small area ($<$ 100 pixels) and with too low mean area ($<$ 10.0 in the background subtracted image) are also removed.

(2) **Building the neighbourhood graph**: Next, the method estimates, which cells are neighbours, i.e. for which cells there is direct contact. If there is a channel with a membrane marker, this can be directly measured (or estimated with high confidence). Here, the contact is inferred from the nuclear signal alone, based on a Voronoï tessellation approach. The Voronoï tessellation partitions the image with respect to a set of seeds in such a way that each point in the space is assigned to the seed to which it is closest. Here, the nuclei are taken as seeds. As there might be points in the image which do not belong to any cell, image points with a distance larger than a user defined value (typically 100 pixels) are discarded, and potential cell contacts beyond this limit are not considered. Based on this method, a neighbourhood graph is built, where the cell nuclei are the nodes and there is an edge between any two cells that were found to be in contact. This neighbourhood graph is the input for the next step.

(3)**Selection of cells:** Each node of the graph generated in the previous step gets one of 3 labels: {cluster, forbidden, unprocessed}. At the start of the algorithm, all nuclei are in state "unprocessed". First, one cell needs to be selected. This selection is in principle arbitrary, but it can be advantageous to select the cell, which is closest to the average cell in terms of area, intensity and texture (in order to avoid selecting particularly exceptional cells for the experiment). This is particularly useful if the cluster size is 1, i.e. if no neighbouring cells are illuminated. For cluster sizes larger than 2, it might seem advantageous to choose the cell that has most connections to other cells, in order to increase cluster compactness. Another option is to choose the cell closest to one of the corners. The selected cell is called "cluster seed". Next, the cell cluster is thus initialized with the cluster seed. The next step is to expand this cluster to match the cluster size criterion. For this, a list of candidate cells is established, which is simply achieved by making the union of the neighbours of the newly added cell with the initial candidate list (which is empty before initialization with the cluster seed). Among this candidate list, one cell is chosen that has the smallest maximal distance to any of the cells already in the cluster. This is then iterated until the defined cluster size is reached. This procedure results in particularly compact clusters, with relatively small Euclidean cluster extension in any direction. Each nucleus in the cluster gets the label "cluster". Given this cluster, the cells that are not to be illuminated can be found by the following procedure: first, all cells which are neighbours to the newly defined cluster cells (and not themselves cluster cells) are considered to be "forbidden" and their label is changed accordingly. Next, if the chosen inter-cluster distance is larger than 1, the "forbidden" label is extended to the neighbours of the "forbidden" cells. This procedure is iterated K times (with K the defined inter-cluster distance). Once the "forbidden" labels are set, the new candidates for a cluster seed are simply all nuclei that are next to a "forbidden" label. This procedure is iterated until no more cells can be chosen. The entire algorithm is slightly more complicated (due to the presence of separated connected components, the need to avoid infinite loops, etc.).

The output of *Selective Illumination* is a list of positions in a .stg format that is sent back to the microscope, ordering the FRAP head to turn on the laser only on these specific locations. The position defining a given cell corresponds to the center of its nucleus. These experiments were conducted on an inverted spinning disk microscope running under MetaMorph, using a 405nm laser beam.

12.1.2 Graphical User Interface

In the Graphical User Interface (Figure 5.2), input and output folders can be defined. The software will process all images that are located in the directory tree under the input folder. The output folder may or may not exist (in which case it will be created by the software). The following subfolders are also generated:

- coordinates: will contain the coordinates of the chosen cells
- img_chosen_cells: for each image in the input folder, two result images are generated: one with the only the chosen cells, one with the all cells, where the chosen cells are highlighted by a different grey level.
- img_debug: contains all intermediate images of steps 1 and 2. This is only for debugging and illustration purposes.
- img_graph_overlay: contains all intermediate images of step 3.
- Metamorph: contains the file with the positions that can be loaded into metamorph. This is the main output used in this project.

The algorithm parameters that can be changed by the user are (1) the size of the cell clusters, i.e. the number of neighbour cells that are going to be illuminated, (2) the distance between the cell clusters, i.e. the number of cells separating different clusters. If this value is set to 2 for instance, this means that between any two cells in different clusters, there are at least 2 cells, which are not illuminated. Also, for any two different clusters, there exist 2 cells (one cell per cluster), which are separated by exactly 2 cells which are not illuminated. (3) The pixel size can be changed. This is important for the coordinates exported in the metamorph position file. Care has to be taken to fill the correct number, depending on the microscope system used. Finally, there is an option "Ordered Selection". This concerns the choice of candidates for the clusters. If the option is chosen, the priority essentially depends on the x- and y- coordinates, and privileges the uppermost and leftmost cell, among the candidates. In principle, this setting can result in more clusters to be illuminated, but clusters might be less compact (in case the cluster size > 2) or the choice might fall accidentally to an aberrant cell (such as a dead cell or a mitotic cell). Finally, the user can also define the kind of output.

13 Bibliography

1. Lodish, H., A. Berk, and S. Zipursky, *Molecular Cell biology, 4th edition.* 2000.
2. Bissell, M.J. and W.C. Hines, *Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression.* Nature medicine, 2011. **17**(3): p. 320-329.
3. Schwann, T. and M. Schleiden, *Microscopic Investigations on the Accordance in the Structure and Growth of Plants and Animals.* London: Syndenham Society, 1847.
4. Regenmortel, M.H.V.V., *Reductionism and complexity in molecular biology.* EMBO Reports, 2004. **5**(11): p. 1016-1020.
5. Waliszewski, P., M. Molski, and J. Konarski, *On the holistic approach in cellular and cancer biology: nonlinearity, complexity, and quasi-determinism of the dynamic cellular network.* Journal of surgical oncology, 1998. **68**(2): p. 70-78.
6. Weinberg, R.A., *Coming full circle—from endless complexity to simplicity and back again.* Cell, 2014. **157**(1): p. 267-271.
7. Kienle, G. and H. Kiene, *From Reductionism to Holism: Systems-oriented Approaches in Cancer Research.* Global Advances in Health and Medicine, 2012. **1**(5): p. 68-77.
8. Elasser, W.M., *Reflections on a theory of organisms: Holism in biology.* 1998.
9. Radisky, D.M., J; Bissell, MJ, *Order and dissorder: the role of extracellular matrix in epithelial cancer.* Cancer Invest, 2002. **20**(1): p. 139-53.
10. Macara, Ian G., et al., *Epithelial Homeostasis.* Current Biology, 2014. **24**(17): p. R815-R825.
11. Lee, H.O. and C. Norden, *Mechanisms controlling arrangements and movements of nuclei in pseudostratified epithelia.* Trends in Cell Biology, 2013. **23**(3): p. 141-150.
12. Rowlands, T.M., et al., *Cadherins: crucial regulators of structure and function in reproductive tissues.* Reviews of reproduction, 2000. **5**(1): p. 53-61.
13. Nelson, W.J., *Regulation of cell–cell adhesion by the cadherin–catenin complex.* Biochemical Society Transactions, 2008. **36**(2): p. 149-155.
14. Yap, A.S., W.M. Brieher, and B.M. Gumbiner, *Molecular and functional analysis of cadherin-based adherens junctions.* Annual review of cell and developmental biology, 1997. **13**(1): p. 119-146.

15. Takeichi, M., *Dynamic contacts: rearranging adherens junctions to drive epithelial remodelling.* Nat Rev Mol Cell Biol, 2014. **15**(6): p. 397-410.
16. Tariq, H., et al., *Cadherin flexibility provides a key difference between desmosomes and adherens junctions.* Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(17): p. 5395-5400.
17. Lie, P.P.Y., C.Y. Cheng, and D.D. Mruk, *Chapter five - The Biology of the Desmosome-Like Junction: A Versatile Anchoring Junction and Signal Transducer in the Seminiferous Epithelium*, in *International Review of Cell and Molecular Biology*, W.J. Kwang, Editor. 2011, Academic Press. p. 223-269.
18. Garrod, D. and M. Chidgey, *Desmosome structure, composition and function.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 2008. **1778**(3): p. 572-587.
19. Green, K.J. and J. Jones, *Desmosomes and hemidesmosomes: structure and function of molecular components.* The FASEB Journal, 1996. **10**(8): p. 871-881.
20. Hartsock, A. and W.J. Nelson, *Adherens and Tight Junctions: Structure, Function and Connections to the Actin Cytoskeleton.* Biochimica et biophysica acta, 2008. **1778**(3): p. 660-669.
21. Zihni, C., et al., *Tight junctions: from simple barriers to multifunctional molecular gates.* Nat Rev Mol Cell Biol, 2016.
22. Anderson, J.M. and C.M. Van Itallie, *Physiology and function of the tight junction.* Cold Spring Harbor perspectives in biology, 2009. **1**(2): p. a002584.
23. Itoh, M., et al., *Direct Binding of Three Tight Junction-Associated Maguks, Zo-1, Zo-2, and Zo-3, with the Cooh Termini of Claudins.* The Journal of Cell Biology, 1999. **147**(6): p. 1351-1363.
24. Meşe, G., G. Richard, and T.W. White, *Gap Junctions: Basic Structure and Function.* Journal of Investigative Dermatology, 2007. **127**(11): p. 2516-2524.
25. Goodenough, D.A. and D.L. Paul, *Gap Junctions.* Cold Spring Harbor Perspectives in Biology, 2009. **1**(1): p. a002576.
26. Lamouille, S., J. Xu, and R. Derynck, *Molecular mechanisms of epithelial–mesenchymal transition.* Nature reviews. Molecular cell biology, 2014. **15**(3): p. 178-196.
27. Kaplan, N.A., X. Liu, and N.S. Tolwinski, *Epithelial Polarity: Interactions Between Junctions and Apical–Basal Machinery.* Genetics, 2009. **183**(3): p. 897-904.
28. Assémat, E., et al., *Polarity complex proteins.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 2008. **1778**(3): p. 614-630.
29. Roignot, J., X. Peng, and K. Mostov, *Polarity in Mammalian Epithelial Morphogenesis.* Cold Spring Harbor Perspectives in Biology, 2013. **5**(2).

30. Bilder, D. and N. Perrimon, *Localization of apical epithelial determinants by the basolateral PDZ protein Scribble.* Nature, 2000. **403**(6770): p. 676-680.
31. Su, W.-H., et al., *Polarity Protein Complex Scribble/Lgl/Dlg and Epithelial Cell Barriers.* Advances in experimental medicine and biology, 2012. **763**: p. 149-170.
32. Turner, J.R., *Intestinal mucosal barrier function in health and disease.* Nature Reviews Immunology, 2009. **9**(11): p. 799-809.
33. Knight, D.A. and S.T. Holgate, *The airway epithelium: structural and functional properties in health and disease.* Respirology, 2003. **8**(4): p. 432-446.
34. Ahola, A., et al., *Video image-based analysis of single human induced pluripotent stem cell derived cardiomyocyte beating dynamics using digital image correlation.* BioMedical Engineering OnLine, 2014. **13**: p. 39-39.
35. Dotti, C.G., C.A. Sullivan, and G.A. Banker, *The establishment of polarity by hippocampal neurons in culture.* The Journal of neuroscience, 1988. **8**(4): p. 1454-1468.
36. Trelstad, R.L., E.D. Hay, and J.-P. Revel, *Cell contact during early morphogenesis in the chick embryo.* Developmental biology, 1967. **16**(1): p. 78-106.
37. Hay, E., *An overview of epithelio-mesenchymal transformation.* Cells Tissues Organs, 1995. **154**(1): p. 8-20.
38. Hay, E.D., *Organization and fine structure of epithelium and mesenchyme in the developing chick embryo.* Epithelial-mesenchymal interactions, 1968. **2**: p. 31-35.
39. Kalluri, R., *EMT: When epithelial cells decide to become mesenchymal-like cells.* The Journal of Clinical Investigation, 2009. **119**(6): p. 1417-1419.
40. Micalizzi, D.S., S.M. Farabaugh, and H.L. Ford, *Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression.* Journal of mammary gland biology and neoplasia, 2010. **15**(2): p. 117-134.
41. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* The Journal of Clinical Investigation, 2009. **119**(6): p. 1420-1428.
42. Zeisberg, M. and E.G. Neilson, *Biomarkers for epithelial-mesenchymal transitions.* The Journal of Clinical Investigation, 2009. **119**(6): p. 1429-1437.
43. Acloque, H., et al., *Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease.* The Journal of Clinical Investigation, 2009. **119**(6): p. 1438-1449.

44. Ferrer-Vaquer, A., M. Viotti, and A.-K. Hadjantonakis, *Transitions between epithelial and mesenchymal states and the morphogenesis of the early mouse embryo.* Cell adhesion & migration, 2010. **4**(3): p. 447-457.
45. Yang, J. and R.A. Weinberg, *Epithelial-Mesenchymal Transition: At the Crossroads of Development and Tumor Metastasis.* Developmental Cell, 2008. **14**(6): p. 818-829.
46. Grant, C.M. and N. Kyprianou, *Epithelial mesenchymal transition (EMT) in prostate growth and tumor progression.* Translational andrology and urology, 2013. **2**(3): p. 202.
47. Willis, B.C. and Z. Borok, *TGF-β-induced EMT: mechanisms and implications for fibrotic lung disease.* American Journal of Physiology-Lung Cellular and Molecular Physiology, 2007. **293**(3): p. L525-L534.
48. Moustakas, A. and C.-H. Heldin, *Induction of epithelial–mesenchymal transition by transforming growth factor β.* Seminars in Cancer biology, 2012. **22**(5–6): p. 446-454.
49. Xiao, D. and J. He, *Epithelial mesenchymal transition and lung cancer.* Journal of Thoracic Disease, 2011. **2**(3): p. 154-159.
50. Xu, J., S. Lamouille, and R. Derynck, *TGF-/beta/-induced epithelial to mesenchymal transition.* Cell Res, 2009. **19**(2): p. 156-172.
51. May, C., et al., *Epithelial-mesenchymal transition and cancer stem cells: a dangerously dynamic duo in breast cancer progression.* Breast Cancer Research, 2011. **13**(1): p. 202.
52. Peinado, H., D. Olmeda, and A. Cano, *Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?* Nat Rev Cancer, 2007. **7**.
53. Batlle, E., et al., *The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells.* Nat Cell Biol, 2000. **2**(2): p. 84-89.
54. Barrallo-Gimeno, A. and M.A. Nieto, *The Snail genes as inducers of cell movement and survival: implications in development and cancer.* Development, 2005. **132**(14): p. 3151-3161.
55. Muqbil, I., et al., *Snail nuclear transport: The gateways regulating epithelial-to-mesenchymal transition?* Seminars in Cancer biology, 2014. **27**(0): p. 39-45.
56. Lee, J., et al., *The epithelial-mesenchymal transition: new insights in signaling, development, and disease.* J Cell Biol, 2006. **172**: p. 973 - 981.
57. De Herreros, A.G., et al., *Snail family regulation and epithelial mesenchymal transitions in breast cancer progression.* Journal of Mammary Gland Biology and Neoplasia, 2010. **15**(2): p. 135-147.
58. Vandewalle, C., F. Van Roy, and G. Berx, *The role of the ZEB family of transcription factors in development and disease.* Cellular and molecular life sciences, 2009. **66**(5): p. 773-787.

59. Gregory, P., et al., *The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1.* Nat Cell Biol, 2008. **10**: p. 593 - 601.
60. Fernandez-Sanchez, M.E., et al., *Mechanical induction of the tumorigenic [bgr]-catenin pathway by tumour growth pressure.* Nature, 2015. **523**(7558): p. 92-95.
61. Takkunen, M., et al., *Snail-dependent and-independent epithelial-mesenchymal transition in oral squamous carcinoma cells.* Journal of Histochemistry & Cytochemistry, 2006. **54**(11): p. 1263-1275.
62. Capaldo, C.T. and I.G. Macara, *Depletion of E-Cadherin Disrupts Establishment but Not Maintenance of Cell Junctions in Madin-Darby Canine Kidney Epithelial Cells.* Molecular Biology of the Cell, 2007. **18**(1): p. 189-200.
63. Hollestelle, A., et al., *Loss of E-cadherin is not a necessity for epithelial to mesenchymal transition in human breast cancer.* Breast Cancer Research and Treatment, 2013. **138**(1): p. 47-57.
64. Liu, X., et al., *Loss of E-cadherin and epithelial to mesenchymal transition is not required for cell motility in tissues or for metastasis.* Tissue Barriers, 2014. **2**(4): p. e969112.
65. Wheelock, M.J., et al., *Cadherin switching.* Journal of Cell Science, 2008. **121**(6): p. 727-735.
66. de Beco, S., *New insights into the regulation of E-cadherin distribution by endocytosis.* Int Rev Cell Mol Biol, 2012. **2**: p. 63-108.
67. Guillot, C. and T. Lecuit, *Mechanics of Epithelial Tissue Homeostasis and Morphogenesis.* Science, 2013. **340**(6137): p. 1185-1189.
68. Gjorevski, N., E. Boghaert, and C. Nelson, *Regulation of Epithelial-Mesenchymal Transition by Transmission of Mechanical Stress through Epithelial Tissues.* Cancer Microenvironment, 2012. **5**(1): p. 29-38.
69. Thiery, J.P., *Epithelial-mesenchymal transitions in tumour progression.* Nat Rev Cancer, 2002. **2**.
70. Braun, S. and B. Naume, *Circulating and Disseminated Tumor Cells.* Journal of Clinical Oncology, 2005. **23**(8): p. 1623-1626.
71. Tsai, J.H. and J. Yang, *Epithelial–mesenchymal plasticity in carcinoma metastasis.* Genes & Development, 2013. **27**(20): p. 2192-2206.
72. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**.
73. Christiansen, J.J. and A.K. Rajasekaran, *Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis.* Cancer research, 2006. **66**(17): p. 8319-8326.
74. Gupta, G.P. and J. Massagué, *Cancer metastasis: building a framework.* Cell, 2006. **127**(4): p. 679-695.

75. Jolly, M.K., et al., *Implications of the hybrid epithelial/mesenchymal phenotype in metastasis.* Frontiers in Oncology, 2015. **5**.
76. Yu, M., et al., *Circulating Breast Tumor Cells Exhibit Dynamic Changes in Epithelial and Mesenchymal Composition.* Science, 2013. **339**(6119): p. 580-584.
77. Jolly, M.K., et al., *Towards elucidating the connection between epithelial–mesenchymal transitions and stemness.* Vol. 11. 2014.
78. Savagner, P., *The epithelial–mesenchymal transition (EMT) phenomenon.* Annals of Oncology, 2010. **21**(suppl 7): p. vii89-vii92.
79. Fabregat, I., A. Malfettone, and J. Soukupova, *New Insights into the Crossroads between EMT and Stemness in the Context of Cancer.* Journal of clinical medicine, 2016. **5**(3): p. 37.
80. Bonnet, D. and J.E. Dick, *Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell.* Nat Med, 1997. **3**.
81. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells.* Cell, 2008. **133**(4): p. 704-15.
82. Flemming, A., *Cancer stem cells: Targeting the root of cancer relapse.* Nat Rev Drug Discov, 2015. **14**(3): p. 165-165.
83. Rubin, *The suppression of morphological alterations in cells infected with Rous sarcoma virus.* Virology, 1960. **12**: p. 14-31.
84. Rubin, H., *What keeps cells in tissues behaving normally in the face of myriad mutations?* BioEssays, 2006. **28**(5): p. 515-524.
85. Stoker, *Conditions affecting transformation by polyoma virus.* Cold Spring Harbor Quant Bio, 1962. **27**: p. 375-86.
86. STOKER, M.G.P., M. SHEARER, and C. O'NEILL, *Growth Inhibition of Polyoma-Transformed Cells by Contact with Static Normal Fibroblasts.* Journal of Cell Science, 1966. **1**(3): p. 297-310.
87. Rubin, H., *Ordered Heterogeneity and Its Decline in Cancer and Aging*, in *Advances in Cancer Research*, F.V.W. George and K. George, Editors. 2007, Academic Press. p. 117-147.
88. Bosch, F.X., et al., *The causal relation between human papillomavirus and cervical cancer.* Journal of Clinical Pathology, 2002. **55**(4): p. 244-265.
89. Syrjänen, K.J., *Spontaneous evolution of intraepithelial lesions according to the grade and type of the implicated human papillomavirus (HPV).* European Journal of Obstetrics & Gynecology and Reproductive Biology, 1996. **65**(1): p. 45-53.
90. Herfs, M. and C.P. Crum, *Laboratory Management of Cervical Intraepithelial Neoplasia: Proposing a new Paradigm.* Adv Anat Pathol, 2013.
91. Marusyk, A. and K. Polyak, *Tumor heterogeneity: causes and consequences.* Biochimica et biophysica acta, 2010. **1805**(1): p. 105.

92. Wagstaff, L., G. Kolahgar, and E. Piddini, *Competitive cell interactions in cancer: a cellular tug of war.* Trends in Cell Biology, 2013. **23**(4): p. 160-167.
93. Baker, N.E., *Cell competition.* Current Biology, 2011. **21**(1): p. R11-R15.
94. Vincent, J.-P., et al., *Steep Differences in Wingless Signaling Trigger Myc-Independent Competitive Cell Interactions.* Developmental Cell, 2011. **21**(2): p. 366-374.
95. Bondar, T. and R. Medzhitov, *p53-mediated hematopoietic stem and progenitor cell competition.* Cell Stem Cell, 2010. **6**(4): p. 309-322.
96. Marusyk, A., et al., *Irradiation Selects for p53-Deficient Hematopoietic Progenitors.* PLoS Biology, 2010. **8**(3): p. e1000324.
97. Vidal, M., D.E. Larson, and R.L. Cagan, *Csk-Deficient Boundary Cells Are Eliminated from Normal Drosophila Epithelia by Exclusion, Migration, and Apoptosis.* Developmental Cell, 2006. **10**(1): p. 33-44.
98. Kajita, M., et al., *Characterization of the interface between normal and transformed epithelial cells.* Nat Cell Biol, 2010. **11**(4): p. 460-467.
99. Davies, P.C.W., L. Demetrius, and J.A. Tuszynski, *Cancer as a dynamical phase transition.* Theoretical Biology & Medical Modelling, 2011. **8**: p. 30-30.
100. Lyons, J.G., et al., *Snail Up-regulates Proinflammatory Mediators and Inhibits Differentiation in Oral Keratinocytes.* Cancer Research, 2008. **68**(12): p. 4525-4530.
101. Izeddin, I., et al., *Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus*, ed. R.H. Singer. 2014.
102. Récamier, V., et al., *Single cell correlation fractal dimension of chromatin: A framework to interpret 3D single molecule super-resolution.* Nucleus, 2014. **5**(1): p. 75-84.
103. Sinha, D.K., et al., *Photocontrol of Protein Activity in Cultured Cells and Zebrafish with One- and Two-Photon Illumination.* ChemBioChem, 2010. **11**(5): p. 653-663.
104. Medici, D., E.D. Hay, and B.R. Olsen, *Snail and Slug Promote Epithelial-Mesenchymal Transition through β-Catenin–T-Cell Factor-4-dependent Expression of Transforming Growth Factor-β3.* Molecular Biology of the Cell, 2008. **19**(11): p. 4875-4887.
105. P Leroy, K.M., *Slug is required for cell survival during partial EMT of HGH induced tubulogenesis.* Molecular Biology of the Cell, 2007. **18**: p. 1943-1957.
106. Brinkhof, B., et al., *Development and evaluation of canine reference genes for accurate quantification of gene expression.* Analytical Biochemistry, 2006. **356**(1): p. 36-43.

107. Yuan, J.S., et al., *Statistical analysis of real-time PCR data.* BMC bioinformatics, 2006. **7**(1): p. 1.
108. Van der Walt, S., et al., *scikit-image: image processing in Python.* PeerJ, 2014. **2**: p. e453.
109. Cohen, D. and A. Müsch, *Apical surface formation in MDCK cells: regulation by the serine/threonine kinase EMK1.* Methods, 2003. **30**(3): p. 269-276.
110. de Beco, S., et al., *Endocytosis is required for E-cadherin redistribution at mature adherens junctions.* Proceedings of the National Academy of Sciences, 2009. **106**(17): p. 7010-7015.
111. Cieply, B., et al., *Suppression of the epithelial-mesenchymal transition by grainyhead-like-2.* Cancer Res, 2012. **72**(9): p. 2440-53.
112. Tran, D.D., et al., *Temporal and Spatial Cooperation of Snail1 and Twist1 during Epithelial–Mesenchymal Transition Predicts for Human Breast Cancer Recurrence.* American Association for Cancer Research, 2011. **9**(12): p. 1644-1657.
113. Dang, H., et al., *Snail1 induces epithelial-to-mesenchymal transition and tumor initiating stem cell characteristics.* BMC Cancer, 2011. **11**(1): p. 1-13.
114. Naber, H.P.H., et al., *Snail and Slug, key regulators of TGF-β-induced EMT, are sufficient for the induction of single-cell invasion.* Biochemical and Biophysical Research Communications, 2013. **435**(1): p. 58-63.
115. Brocard, J., et al., *Spatio-temporally controlled site-specific somatic mutagenesis in the mouse.* Proc Natl Acad Sci U S A, 1997. **94**(26): p. 14559-63.
116. Feil, R., et al., *Ligand-activated site-specific recombination in mice.* Proc Natl Acad Sci U S A, 1996. **93**(20): p. 10887-90.
117. Domínguez, D., et al., *Phosphorylation Regulates the Subcellular Location and Activity of the Snail Transcriptional Repressor.* Mol Cell Biol, 2003. **23**(14): p. 5078-89.
118. Picard, D., et al., *Signal transduction by steroid hormones: nuclear localization is differentially regulated in estrogen and glucocorticoid receptors.* Cell regulation, 1990. **1**(3): p. 291-299.
119. Reffay, M., et al., *Interplay of RhoA and mechanical forces in collective cell migration driven by leader cells.* Nat Cell Biol, 2014. **16**(3): p. 217-223.
120. Li, H., et al., *Snail Involves in the Transforming Growth Factor β1-Mediated Epithelial-Mesenchymal Transition of Retinal Pigment Epithelial Cells.* PLoS ONE, 2011. **6**(8): p. e23322.
121. Xie, L., et al., *Activation of the Erk Pathway Is Required for TGF-β1-Induced EMT In Vitro.* Neoplasia, 2004. **6**(5): p. 603-610.
122. Grände, M., et al., *Transforming growth factor-β and epidermal growth factor synergistically stimulate epithelial to mesenchymal transition (EMT)*

through a MEK-dependent mechanism in primary cultured pig thyrocytes. Journal of Cell Science, 2002. **115**(22): p. 4227-4236.

123. O'Connor, J.W., et al., *Cell-cell contact and matrix adhesion promote aSMA expression during TGFβ1-induced epithelial-myofibroblast transition via Notch and MRTF-A.* Scientific Reports, 2016. **6**: p. 26226.
124. Fan, P., et al., *Identification of gene regulation patterns underlying both E(2)- and tamoxifen-stimulated cell growth through global gene expression profiling in breast cancer cells.* European journal of cancer (Oxford, England : 1990), 2014. **50**(16): p. 2877-2886.
125. Nicolas, F.J., et al., *Epithelial mesenchymal transition in MDCK cells is accompanied by downregulation of Smad3 expression leading to resistance toTGFβ-induced growth arrest.* Journal of Biological Chemistry, 2002.
126. *Method of the Year 2010.* Nat Meth, 2011. **8**(1): p. 1-1.
127. Deisseroth, K., *Optogenetics.* Nat Meth, 2011. **8**(1): p. 26-29.
128. Fenno, L., O. Yizhar, and K. Deisseroth, *The development and application of optogenetics.* Neuroscience, 2011. **34**(1): p. 389.
129. Boyden, E.S., et al., *Millisecond-timescale, genetically targeted optical control of neural activity.* Nat Neurosci, 2005. **8**(9): p. 1263-1268.
130. Tischer, D. and O.D. Weiner, *Illuminating cell signalling with optogenetic tools.* Nat Rev Mol Cell Biol, 2014. **15**(8): p. 551-558.
131. Pathak, G.P., J.D. Vrana, and C.L. Tucker, *Optogenetic Control of Cell Function Using Engineered Photoreceptors.* Biology of the cell / under the auspices of the European Cell Biology Organization, 2013. **105**(2): p. 59-72.
132. Kim, B. and M.Z. Lin, *Optobiology: Optical control of biological processes via protein engineering.* Biochemical Society transactions, 2013. **41**(5): p. 1183-1188.
133. Ádám, É., et al., *Optogenetics: past, present and future.* Acta Biologica Szegediensis, 2015. **59**(Supl 1): p. 105-119.
134. Andresen, M., et al., *Structural basis for reversible photoswitching in Dronpa.* Proceedings of the National Academy of Sciences, 2007. **104**(32): p. 13005-13009.
135. Collares-Buzato, C., et al., *Co-culture of two MDCK strains with distinct junctional protein expression: a model for intercellular junction rearrangement and cell sorting.* Cell and tissue research, 1998. **291**(2): p. 267-276.
136. Aujard, I., et al., *o-Nitrobenzyl Photolabile Protecting Groups with Red-Shifted Absorption: Syntheses and Uncaging Cross-Sections for One- and Two-Photon Excitation.* Chemistry – A European Journal, 2006. **12**(26): p. 6865-6879.

137. Neveu, P., et al., *A Caged Retinoic Acid for One- and Two-Photon Excitation in Zebrafish Embryos.* Angewandte Chemie International Edition, 2008. **47**(20): p. 3744-3746.
138. Rubin, H., *Cell–cell contact interactions conditionally determine suppression and selection of the neoplastic phenotype.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(17): p. 6215-6221.
139. Gurskaya, N.G., et al., *Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light.* Nature biotechnology, 2006. **24**(4): p. 461-465.
140. Chudakov, D.M., S. Lukyanov, and K.A. Lukyanov, *Tracking intracellular protein movements using photoswitchable fluorescent proteins PS-CFP2 and Dendra2.* Nature protocols, 2007. **2**(8): p. 2024-2032.
141. Zahid, M., et al., *Holographic Photolysis for Multiple Cell Stimulation in Mouse Hippocampal Slices.* PLoS ONE, 2010. **5**(2): p. e9431.
142. Ronzitti, E., et al., *LCoS nematic SLM characterization and modeling for diffraction efficiency optimization, zero and ghost orders suppression.* Optics Express, 2012. **20**(16): p. 17843-17855.
143. Papagiakoumou, E., et al., *Patterned two-photon illumination by spatiotemporal shaping of ultrashort pulses.* Optics Express, 2008. **16**(26): p. 22039-22047.
144. Liang, J., et al., *Suppression of the zero-order diffracted beam from a pixelated spatial light modulator by phase compression.* Applied optics, 2012. **51**(16): p. 3294-3304.
145. Jesacher, A., S. Bernet, and M. Ritsch-Marte, *Broadband suppression of the zero diffraction order of an SLM using its extended phase modulation range.* Optics express, 2014. **22**(14): p. 17590-17599.
146. Palima, D. and V.R. Daria, *Holographic projection of arbitrary light patterns with a suppressed zero-order beam.* Applied optics, 2007. **46**(20): p. 4197-4201.
147. Carpi, N., et al., *Micropatterning on glass with deep UV.* 2011.
148. Sternberg, P.W. and H.R. Horvitz, *Pattern formation during vulval development in C. elegans.* Cell, 1986. **44**(5): p. 761-772.
149. Hill, R.J. and P.W. Sternberg, *Cell fate patterning during C. elegans vulval development.* Development, 1993. **119**(Supplement): p. 9-18.
150. Beira, J.V. and R. Paro, *The legacy of Drosophila imaginal discs.* Chromosoma, 2016: p. 1-20.
151. Gomez, E.W., et al., *Tissue geometry patterns epithelial–mesenchymal transition via intercellular mechanotransduction.* Journal of Cellular Biochemistry, 2010. **110**(1): p. 44-51.

152. Garcia, S., et al., *Physics of active jamming during collective cellular motion in a monolayer.* Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(50): p. 15314-15319.
153. Satir, P., L.B. Pedersen, and S.T. Christensen, *The primary cilium at a glance.* Journal of Cell Science, 2010. **123**(4): p. 499-503.
154. Rozycki, M., et al., *The fate of the primary cilium during myofibroblast transition.* Molecular Biology of the Cell, 2014. **25**(5): p. 643-657.
155. Clement, C.A., et al., *TGF-β signaling is associated with endocytosis at the pocket region of the primary cilium.* Cell reports, 2013. **3**(6): p. 1806-1814.
156. Senoo-Matsuda, N. and L.A. Johnston, *Soluble factors mediate competitive and cooperative interactions between cells expressing different levels of Drosophila Myc.* Proceedings of the National Academy of Sciences, 2007. **104**(47): p. 18543-18548.
157. Vivarelli, S., L. Wagstaff, and E. Piddini, *Cell wars: regulation of cell survival and proliferation by cell competition.* Essays in biochemistry, 2012. **53**: p. 69-82.
158. Kosaka, N., et al., *Competitive interactions of cancer cells and normal cells via secretory microRNAs.* Journal of Biological Chemistry, 2012. **287**(2): p. 1397-1405.
159. Zhou, W., et al., *Cancer-secreted miR-105 destroys vascular endothelial barriers to promote metastasis.* Cancer cell, 2014. **25**(4): p. 501-515.
160. Roskelley CD, B.M., *The dominance of the microenvironment in breast and ovarian cancer.* Semin Cancer Biol, 2002. **12**(2): p. 97-104.
161. Leight, J.L., et al., *Matrix rigidity regulates a switch between TGF-β1–induced apoptosis and epithelial–mesenchymal transition.* Molecular Biology of the Cell, 2012. **23**(5): p. 781-791.
162. Toyli, M.R.-K.L.C.J.V.J.E.S., *Different responses in transformation of MDCK cells in 2D and 3D culture by v-Src as revealed by microarray techniques, RT-PCR and functional assays.* Lab Invest, 2010. **90**(6): p. 915-28.
163. Lucas, A.M.S., *Impact of the 3D microenvironment on phenotype, gene expression and EGFR inhibition of colorectal cancer cell lines.* PloS one, 2013. **8**(3).
164. O'Brien, L.E., et al., *Morphological and Biochemical Analysis of Rac1 in Three - Dimensional Epithelial Cell Cultures*, in *Methods in Enzymology.* 2006, Academic Press. p. 676-691.
165. Elia, N. and J. Lippincott-Schwartz, *Culturing Three Dimensional MDCK cells for Analyzing Intracellular Dynamics.* Current protocols in cell biology / editorial board, Juan S. Bonifacino ... [et al.], 2009.: p. Unit-4.22.
166. Temin, R., *Characteristics of an assay for Rous sarcoma virus and Rous sarcoma cells in tissue culture.* Virology, 1958. **6**: p. 669-88.

Summary

Tissues and organisms are built from cells in which important phenotype decisions are made: division, differentiation, apoptosis, and transformation. Cell biology has strongly focused on deciphering the internal molecular determinants of these decisions, but external information originating from intercellular interactions are key elements to coordinate multicellular physiology.

The extent to which internal determinants dominate over external determinants or *vice versa*, is an essential feature of the sociology of cell communities, with possibly strong individualistic situations, or dominant collective effect.

The present work was designed to set-up a method for assessing the relative contribution of internal vs. external determinant, by opposing these two classes of inputs. This is achieved by challenging the collective stability of an *in vitro* epithelium using the heterogeneous induction of the epithelial-to-mesenchymal transition (EMT) via the photoactivation of Snail1. The key results show that the transcriptional response of EMT-induced cells depends on the presence of non-induced cells in the culture. Conversely non-induced cells respond to the presence of induced cells. These mutual control effects lead to the notion that the geometry underlying the distribution of a given molecular cause strongly influences its consequence. Our work opens new perspectives for studying the sociology of heterogeneous cell communities, and better understand important phenomena such as phenotype suppression and or the onset of carcinogenesis.

Résumé

L'ensemble des tissus et des organismes vivants sont constitués de cellules dans lesquelles un certain nombre de décisions phénotypiques sont prises : division, différentiation, apoptose ou encore transformation. La biologie cellulaire s'est principalement concentrée sur la compréhension des déterminants moléculaires internes de ces décisions, mais il est important de considérer aussi l'existence de déterminants externes provenant des interactions intercellulaires qui sont essentielles à l'émergence de systèmes multicellulaires coordonnés.

La compétition entre les déterminants internes et les déterminants externes est un aspect fondamental de la sociologie des communautés cellulaire menant à de possibles situations hautement individualisées ou, au contraire, à un effet collectif dominant.

Ce travail de thèse a eu pour but de mettre en place une méthode permettant de mesurer la contribution relative de ces deux types de déterminants en les mettant en opposition. Pour cela, la stabilité collective d'un épithélium *in vitro* a été mise à l'épreuve grâce à l'induction hétérogène de la transition épithelio-mesenchymateuse (EMT) par le biais de la photoactivation du facteur de transcription Snail1. Les résultats principaux montrent que la réponse transcriptionelle de cellules induites à l'EMT dépend de la présence, ou non, de cellules avoisinantes non-induites. De la même manière, les cellules non-induites répondent de façon transcriptionelle à la présence de cellules induites. Ces effets de control mutuels introduisent la notion que la géométrie de la distribution d'une cause moléculaire donnée peut influencer la conséquence de cette même cause. Notre travail ouvre de nouvelle possibilités pour l'étude de la sociologie de communautés cellulaires hétérogènes, et une meilleure compréhension de phénomènes importants tel la suppression phénotypique ou encore les premiers instants de la carcinogenèse.