Principles of Development

– An inquisitive journey into embryology –

Muhammed Fethullah Simsek

Edited by: Andrea Alexei

This e-book is accessible as an Open Educational Resource (OER) through: https://ecampusontario.pressbooks.pub/principlesofdev/

and can be cited using the digital object identifier (DOI): 10.5281/zenodo.22545177.

Contents

Foreword

This little notebook came out of a need. The need to cover key concepts of embryology and principles of development in an understandable but detailed enough pace for advanced undergraduate students. Apprentices of those notes are expected to come with a battery of fundamental science knowledge and basic exposure to introductory cell biology and molecular genetics. However, a keen interest to ask questions and seek knowledge may suffice not to lose the thread of discussion. The focus of narrative is on the key processes of the development instead of dry terminology or descriptive storytelling. Although yet immature to talk about any universal laws of nature, our understanding of embryogenesis has developed significantly with the molecular biochemistry revolution in last quarter of 20th century, establishment of model organisms with whole genome sequences in the dawn of the 21st, and advances in microscopy tools and techniques over the last couple decades. Benefiting from many valuable resources and recent publications in the field, I tried to focus on this quantitative understanding throughout the text. Lewis Wolpert's "*Developmental Biology: A Very Short Introduction*" from Oxford Press has been inspiring to draft the first outline. "*Principles of Development*" by Wolpert, Tickle, and Martinez Arias, "*Developmental Biology*" by Gilbert and Barresi, and "*Mechanisms of Morphogenesis*" by Davies has been three key textbook resources along the road. Wherever possible, a not well-trodden path was taken in the transdisciplinary realm across life and physical sciences to satiate the joyful inquiry of 'why'.

The language is intentionally kept colloquial with the raw material originating from lecture recordings. I am grateful to the generous support from MacPherson Institute which enabled me to bring together a brilliant team to sculpt these piecemeal recordings into a coherent structure. Without altruistic efforts of *Andrea Alexei* on typewriting, organising, and editing the notes; and enthusiastic creations of *Henin Ye* for some figures sprinkled throughout, this book would not have been realized. Lastly, I am always grateful and deeply indebted to my spouse, *Fatma*, and young family at various developmental stages, to be accommodating and supporting throughout many edits and their curious feedback along the process.

M.F. Simsek, Burlington, Ontario.

Chapter 1: FROM SO SIMPLE A BEGINNING…

"THERE IS GRANDEUR IN THIS VIEW OF LIFE… WHILST THIS PLANET HAS GONE CYCLING ON ACCORDING TO THE FIXED LAW OF GRAVITY, FROM SO SIMPLE A BEGINNING ENDLESS FORMS MOST BEAUTIFUL AND MOST WONDERFUL HAVE BEEN, AND ARE BEING, EVOLVED."

– C. Darwin, 1859.

Charles Darwin concluded his magnum opus 'On the Origin of Species' with these sentences. Drawing parallels, we can state that endless forms have been and are being **developed** from so simple a beginning, a fertilized single cell, as each multicellular organism gives birth to progenies on 'the pale blue dot' we live in.

"Where do we come from" has been a perplexing question for many since the dawn of humanity. With the advents of technology in 20th century, we first were able to observe this simple beginning of ours under microscopes with *in vitro* fertilization (IVF) of human embryos. How does a single **fertilized** egg give rise to an adult? Does embryonic development have its own fixed laws, as Darwin was mentioning for gravity? More specifically, how do the **differentiated** cells carrying out particular tasks arise? How do those cells **form** the tissues and organs? How is the **pattern** generated? This course is an invitation to embark on a journey of answering those questions and many more to arise along the path.

We will try to understand **'differentiation'** by asking what mechanisms generate the cell diversity while, with rare exceptions, each cell has the same genetic information. We will dive deeper into **'morphogenesis'** (the creation of form) by trying to understand how the cells in our body are organized and coordinated to generate organs. We will discuss **'growth'** from the perspective of how cell divisions and growth are precisely regulated during embryogenesis. We will make sense of **'reproduction'** by answering how the specialized cells of the germ line are set aside. We will get some clues on **'regeneration'** by discussing what regulates cellular and developmental potential. Frequently, our discussions will lead to certain implications about **'speciation'**, *i.e.,* how changes in development over time can generate new body forms. We will further touch **'environment'** via how the environmental factors impact embryonic development and beyond.

Development is a process of becoming instead of being. Coherently, our focus in this journey will be more on what builds the embryo rather than what an embryo is. Builders of the embryo are the cells which act through time and in space. We can speak of four main cellular activities relevant to development.

First one is **'multiplication'**. That is a cell making a greater number of cells, which can manifest as proliferation or cleavage. Second category is **'movement'**. Cells display motility, migration, convergent extension, invagination, ingression, delamination events over the course of embryonic development. Third is **'change in character'**. This involves induction and commitment, specification, and differentiation. Last category of cellular activities is **'signalling'**. Cells as the builders of an embryo will frequently need to communicate with each other via paracrine and juxtacrine paths, or specialized extensions *e.g.,* cytonemes and cilia.

Chapter 2: BASIC PRINCIPLES OF EMBRYOGENESIS

"OUR REAL TEACHER HAS BEEN AND STILL IS THE EMBRYO, WHO IS, INCIDENTALLY, THE ONLY TEACHER WHO IS ALWAYS RIGHT."

– V. Hamburger (1900-2001).

In the first chapter, we concluded our discussion on the builders of the embryo, which are the cells. The concept that cells being the builders of the embryo may seem like a paradoxical chicken-and-egg problem. In the end, we know that the embryo makes all the different cells, but also that the cells themselves are responsible for forming the embryo. As we progress through the course, I hope this apparent contradiction will be relieved and what we mean by 'cells are building the embryo' will become clearer.

What can a cell do?

Before diving deep into the complexity of embryogenesis, let us contemplate on things a single cell can do. We will briefly discuss five levels of activities. First off, a single cell can effectively search for and track the sources of sustenance (food) for supporting its life **(Figure 2.1a)**. A classical and well-understood example is chemotaxis, bacteria's search for nutrition. At bacteria's length-scale (~1 μm), water acts as a highly viscous fluid. To swim through is a peculiar task. Each bacterium is equipped with multiple flagella with corkscrew type rotational chirality. Flagella can switch between rotating clockwise or counter. Due to structural chirality, counterclockwise rotations bundle the flagella together whereas clockwise ones spread them apart, resulting in 'long run' or 'tumble at place' modes of motion, respectively. Bacteria's molecular network can react to changes of chemical nutrient signals, *e.g.,* sugar, over time by switching direction of flagellar rotation more or less frequently. More frequent switches enable change of direction by tumbles. Less frequent switches at steeper nutrient changes will result in longer distances spanned, which puts bacteria closer to the centre of nutrient resources.

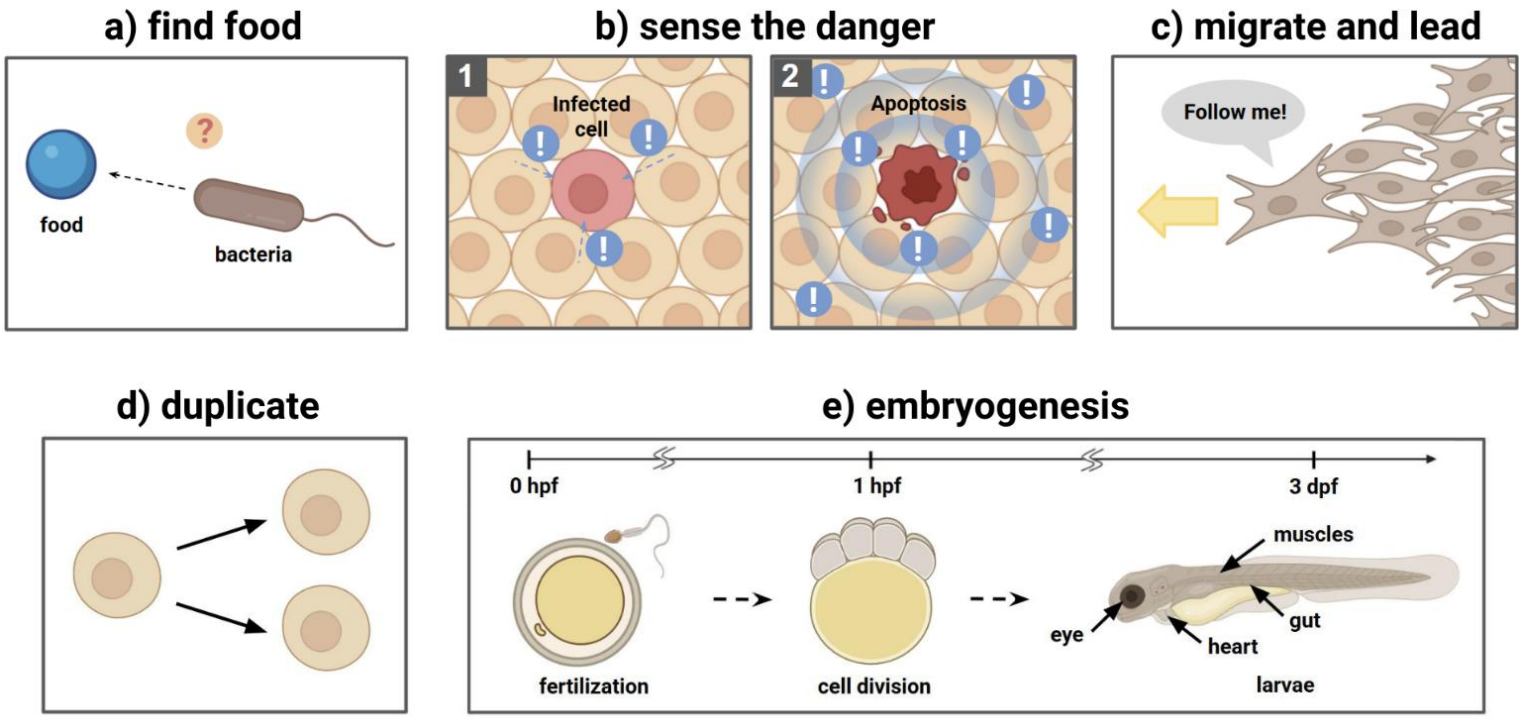

Figure 2.1 | Cells have unique capabilities of measuring diffusive chemicals effectively (**a**), communicating messages across neighbours (**b**), migrating and leading others (**c**), and making copies of themselves (**d**). All those features make them ideal 'builders of the multicellular embryo' (**e**).

Cells can further sense danger and communicate it to others **(Figure 2.1b)**. For instance, when certain cells spread in an epithelium patch are infected with virus and undergo apoptosis, the programmed cell death, this would trigger waves of calcium activity as a communication of danger between cells.

Third, cells can migrate and lead others in migration **(Figure 2.1c)** as observed in the classical example of 'wound healing' in epithelium. When the surface of epithelium is scratched off to create a gap, cells at the

edge will initiate the migration activity by treadmilling their actin cytoskeleton to close the gap and will lead others to follow in action driving **'collective cell migration'**.

At a deeper level, what if one tells you that they devised some sort of robotics or automaton which can 1- trace substances needed for its action, 2- avoid dangers, 3- lead other automatons to move in groups, and 4- further make clones of themselves. This is **'cell division'**, a feat which sounds so mundane to us because of our superficial familiarity with the phenomenon since early education **(Figure 2.1d)**.

Lastly, a single cell, as it keeps making copies of itself and giving rise to 2, 4, 8, 16, 32, 64 and so on daughter cells, can develop into a full functionally specialized multicellular organism: the process we call **'development'**. One can easily say this is the most fascinating aspect of a single cell's capabilities **(Figure 2.1e)**. The collective capacity of a cell's functions allows for the formation of multicellular and developmentally complex life-forms, from the nematode to the zebrafish to humans. For example, division coupled with migration gives rise to the phenomena of gastrulation, where the collective migration mediates the formation of the head-to-tail axis in the body. As it stands, we currently understand that **all the information necessary for embryonic development is contained within the cell, the fertilized egg.** However, this assertion is a significant claim, and the scientific path to it was neither straightforward nor rapid. A brief historical overview of early philosophical perspectives is instructive.

Brief history of embryology

Our overview will begin with **Aristotle** (350 B.C.) who observed embryonic development in birds and fish. Aristotle's world view was that presence of four causes leads something into existence: material, formal, efficient, and final. Let us use the example of a wooden chair. The material cause is the substance, which in the case of the chair would be the wood, while the formal cause is the design of the object. The efficient cause is the effort and actual construction of the chair, and the final cause is the intended purpose, such as sitting. Under the light of this philosophical framework and his observations on chick embryos, Aristotle posited presence of an active and passive faculty within the developing embryo. He thought that a formless mass sculpted into the embryo, which is correct. But Aristotle thought the efficient cause was the male germ which acts on the passive menstrual blood of the female as active faculty, which is quite wrong. This overall framework

of “active” and “passive” faculties would influence other philosophers and early scientists for quite some time. **Galen** (circa 150 A.D.), a founding figure of medicine, proposed that both sexes contribute material, active and passive elements, for the formation of the embryo. However, he believed that one sex’s contribution was greater than the other, which also is inaccurate.

In the 1200s, a more nuanced perspective emerged. An interesting figure, **Ibn an-Nafis**, said that neither of the sexes has the fashioning power, but only the mixing of the two allows for embryonic development. The proposal of ibn an-Nafis was a relatively correct understanding of embryology, ahead of his time. But it is important to remember that Aristotle's observations (primarily of chick embryos) and other early theories mentioned so far were not real ‘science’ as we understand it today. They were more of ideations with philosophical or theological drives. And this idea, which was proposed in 1200s, did not get much traction even by the late 1600s in Europe. For instance, a prominent Renaissance figure **William Harvey**, who was most possibly aware of Ibn an-Nafis’ works, still believed that sperm, upon entering the womb, metamorphosed into an entity that creates the embryo. For a quite long time, we did not understand how the embryo came to be and where the information of its construction came from.

However, in the 18th century things started to change. The advent of microscopy and its increasing use led to the rise of **descriptive embryology**. However, observation alone did not reveal the nuances of embryological development. Marcello Malpighi, who meticulously studied chick embryo development at a microscopic level, still concluded that miniature organs existed within the embryo and simply grew larger. This concept is known as **preformationism**. It was the work of Caspar Friedrich Wolff, a century later, that would refute this theory and ultimately form the foundations for our current understanding. Employing time-lapse microscopy (by the standards of his time), Wolff carefully observed the chick embryo development and demonstrated that organs arise *de novo* from different embryonic structures and tissues.

By the 19th century, descriptive embryology gave rise to the field of **comparative embryology**. The bedrock of modern embryological studies, comparative embryology observes both the differences and similarities of embryonic development between diverse species. Several key figures, including Pander, von Baer, and Kowalevski, made groundbreaking discoveries. These discoveries established fundamental principles

relevant to the developmental processes that guide the formation of all multi-cellular life which we will dig deeper in this chapter.

The fundamentals of embryogenesis

19th and early 20th century observations of comparative embryologists led to certain fundamental rules for embryogenesis which we know hold true to this day. Before diving deeper, I want to note here that the expectation for this course is not to memorize terminology but to critically assess and understand implications. More important than being able to list off definitions is seeing the "big picture" of how life **happens**.

1. A generalizable life cycle

While certain characteristics of embryogenesis may vary, such as the germs layers which you will learn more about in the following section, development for all multicellular forms of life can be unified under certain stages (**Figure 2.2**).

This generalizable life cycle begins with the fusion of male and female germ cells, which are haploid (having single copy of chromosomes): the egg and the sperm. The event of

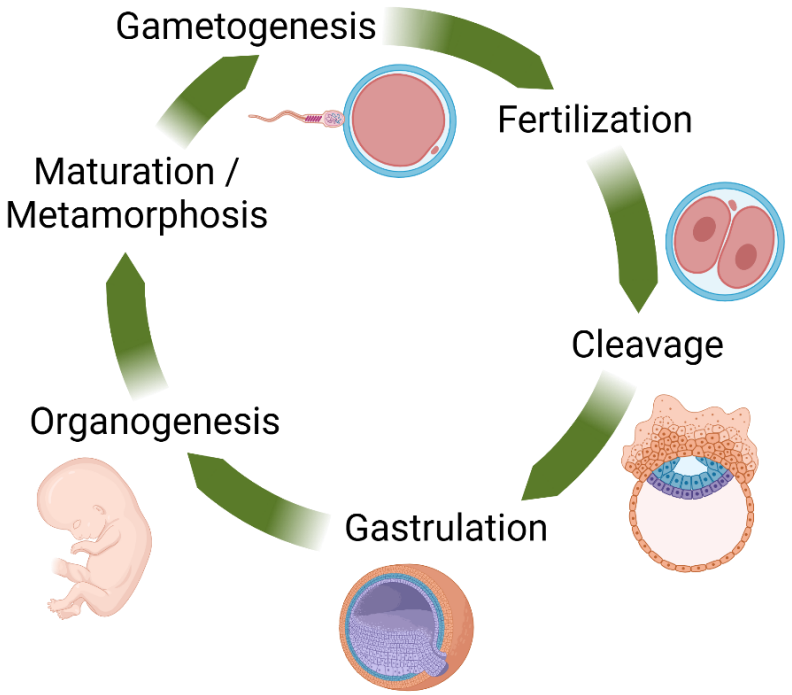

Figure 2.2 | Generalizable life cycle in multicellular organisms.

their fusion is referred to as **fertilization** and gives rise to the diploid **zygote** (genetic material from two sexes is now mixed). The zygote then undergoes a series of cleavage events. Cleavage is a specific type of mitotic cell division where the cells divide rapidly without increasing in size. After cleavage cycles, **blastula**, a multi-cellular collection that shares comparable total volume to the original zygote, forms. At this stage, the cells are relatively undifferentiated.

The cleavage patterns amongst different organisms are remarkably diverse, due to a major influence by the **yolk geometry**. During embryogenesis, the yolk serves only as a nutrient source and does not contribute to animal tissue. In certain animals, such as mammals and nematodes, the yolk is evenly distributed and embryos undergo a cleavage pattern where they divide symmetrically until the cavity of the blastula, the **blastocoel**, forms. Species with uneven yolk distributions can undergo different forms of cleavage. In amphibian development, where yolk is denser on one side of the egg, we see that the side resulting in having smaller cells becomes the animal pole of the embryo. In fish, where the yolk is even more localized to one side, a planar division pattern known as **discoidal cleavage** takes place. Birds also exhibit a discoidal cleavage pattern. They have a massive yolk, and the embryo develops on top of it, encased in a membrane called the vitelline membrane.

Insect embryos present a unique scenario. During cleavage, there is no distinct boundary between cells. The nuclei divide, but the cells share a single cytoplasm, with the yolk located internally, surrounded by the nuclei. Despite the differences, all cleavage patterns share a common feature: the egg material is progressively divided into smaller units as the cleavage progresses.

Once the blastula is formed **gastrulation** takes place. At this stage of development, we observe collective cell migration events that give rise to the **germ layers**. As differentiation leads to more cell types during development, each germ layer gives rise to organ structures during **organogenesis**. Once this process is complete, the offspring will be born either through birth or hatching. The offspring will undergo **maturation or metamorphosis** as the next phase of shared life cycle, giving rise to a sexually mature adult. The mature organism will use the haploid germ cells produced through **gametogenesis** to contribute to the gene pool of next generation, hence completing ‘the life cycle’.

2. Germ layers

As gastrulation separates the blastula into defined domains, the germ layers are developed. Relative to the yolk, the **ectoderm** forms the outer layer, the **mesoderm** the middle layer, and the **endoderm** the inner layer. Besides the **germ cells**, all cell types in an adult organism derive from these layers. However, not all organisms develop all three germ layers. Members of the Cnidaria, Ctenophora, and Porifera phyla are **diploblastic**, meaning they only form an ectoderm and endoderm. All other animals are **triploblastic**, developing the mesoderm.

Cells specified to make up each germ layer give rise to unique organ structures (**Figure 2.3**). The **ectoderm** develops into the outer surface organ, which for humans would be the skin, the central nervous system and, a unique innovation in vertebrates, the neural crest. Neural crest cells are found only in vertebrates and their remarkable ability to differentiate into a wide range of tissues, including skin, muscle, and parts of internal organs, makes them one-of-a-kind. They undergo some interesting cell migration dynamics within the ambient development. For instance, with unique migration paths, neural crest cells

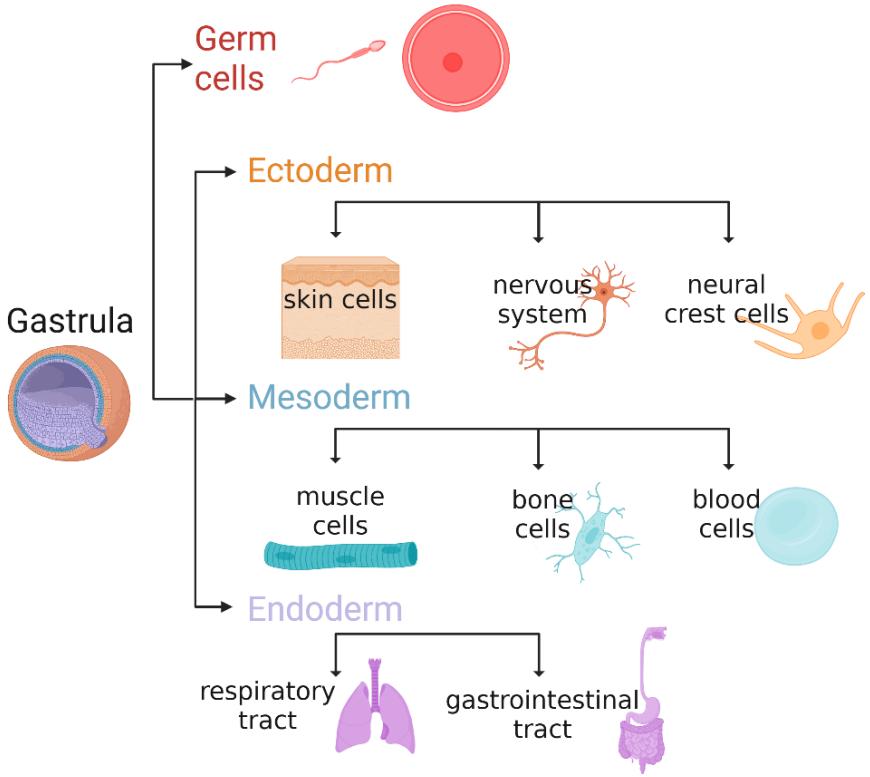

Figure 2.3 | Embryonic germ layers and example animal cells and tissues they contribute to in the adult organism.

sculpted our facial structure when we were an embryo, giving every one of us unique faces. The **mesoderm**, on the other hand, contribute to the dorsal, paraxial, intermediate, lateral, and head structures. This leads to the development of bones, blood vessels, blood cells, and facial muscles in the embryo. The **endoderm** comprises cells which will form mostly inner organs, *e.g.,* the digestive tube, esophagus, and respiratory tract of the offspring (**Figure 2.3**). Altogether, these distinct layers are the first separation of tasks and lead to the development of specialized cellular and organ structures in embryogenesis.

Proliferation comes after rapid rounds of early mitotic divisions (cleavage). During organogenesis, which follows the formation of the distinct germ layers, we observe the cells to mainly proliferate, not cleave. Unlike cleavage, we see an increase in total cytoplasmic content and cell volume during proliferation. With organogenesis, the multicellular volume of the embryo **grows**, which is not the case in the cleavage and the blastula formation.

3. The phylotypic stage

Early comparative embryologists established that embryonic development involves **'a gradual process of specification and divergence of form'**. While the yolk geometry and additional embryonic tissues show greater variability between organisms, the tissues that ultimately gives rise to embryo itself, the **embryo proper**, become quite similarly shaped amongst vertebrates. These most similar stages are referred to as the **phylotypic stage**.

What guides the observed similarities, and when do they take place? During the initial division of the fertilized zygote, there is weaker selective pressure on embryonic form. Owing to this, embryos at very early stages lack much similarity. However, as vertebrates begin forming **somites** during somitogenesis, they enter a crucial developmental funnel that connects gastrulation to organogenesis. This stage, characterized by a particularly strong selective pressure on how the embryo develops before organ specification, exhibits the greatest degree of morphological similarity. Following this stage the selective pressure on the embryonic form weakens, and a greater disparity is seen between developing embryos of different organisms.

The **"hourglass model"** encapsulates these observations. This model posits that the phylotypic stage occurs during mid-embryogenesis. That is to say, the selective pressure during somitogenesis funnels the

developmental processes, which are otherwise distinct in very early and late embryogenesis. While the concept of a phylotypic stage was originally generated due to physical observations of embryological development, we now also consider the molecular signals and patterns of gene and protein expression when determining "similarity" between the embryonic development of diverse organisms.

4. Induction / specification

The observation of embryogenesis alone cannot answer a straightforward but elusive question: what controls this process and where is the information stored? Experiments involving the dissection and isolation of germ layers revealed that they cannot develop independently. These observations led to the concept of induction, where one tissue or cell signals to another, instructing it to differentiate into a specific cell type, and suggested the critical role of intercellular communication during development. The works conducted by Hans Driesch, a prominent figure in the field in late 19th and early 20th centuries, led to the discovery of **regulative development** and identified the importance **cell-to-cell communication** during development (**Figure 2.4**).

Driesch was able to identify the phenomenon of **regulative development** by studying sea urchin embryos at the four-cell stage (**Figure 2.4a**). By removing the fertilization envelope and vigorously shaking, Driesch split the four cells apart without damaging them. Afterwards, he observed their independent development. Remarkably, each cell developed into a complete, albeit smaller, sea urchin. Through this experimental process, one can conclude that all the information for embryonic development is in the fertilized egg; and further understand that not only can a single fertilized cell, but the pieces of that fertilized cell as well, give rise to an embryo.

The importance of **cell-to-cell communication** was realized through Driesch's further work on the eight-cell and sixteen-cell stage sea urchin embryos (**Figure 2.4b**). From four to eight-cell stage, division axis of the cells changes perpendicularly, giving rise to two distinct subgroups (animal vs. vegetal pole). To assess the importance of this relative positioning, Driesch took the embryos and pushed them between two glass slides. The applied pressure disrupted the division pattern and distribution of the embryonic cells. Despite the disruption, embryos developed normally, indicating that the cells could compensate for positional changes. Today, we know that these observations are not universally true for all animals, but that the sea urchin embryo

is more resilient to perturbations. However, Driesch's findings did correctly conclude that cell-to-cell communication influences embryonic development. These discoveries took place in the dawn of 20th century when scientists barely knew the existence of chromosomes and did not know the genetic code. Unaware of the signalling molecules guiding intercellular communications, they understood the cells needed to communicate.

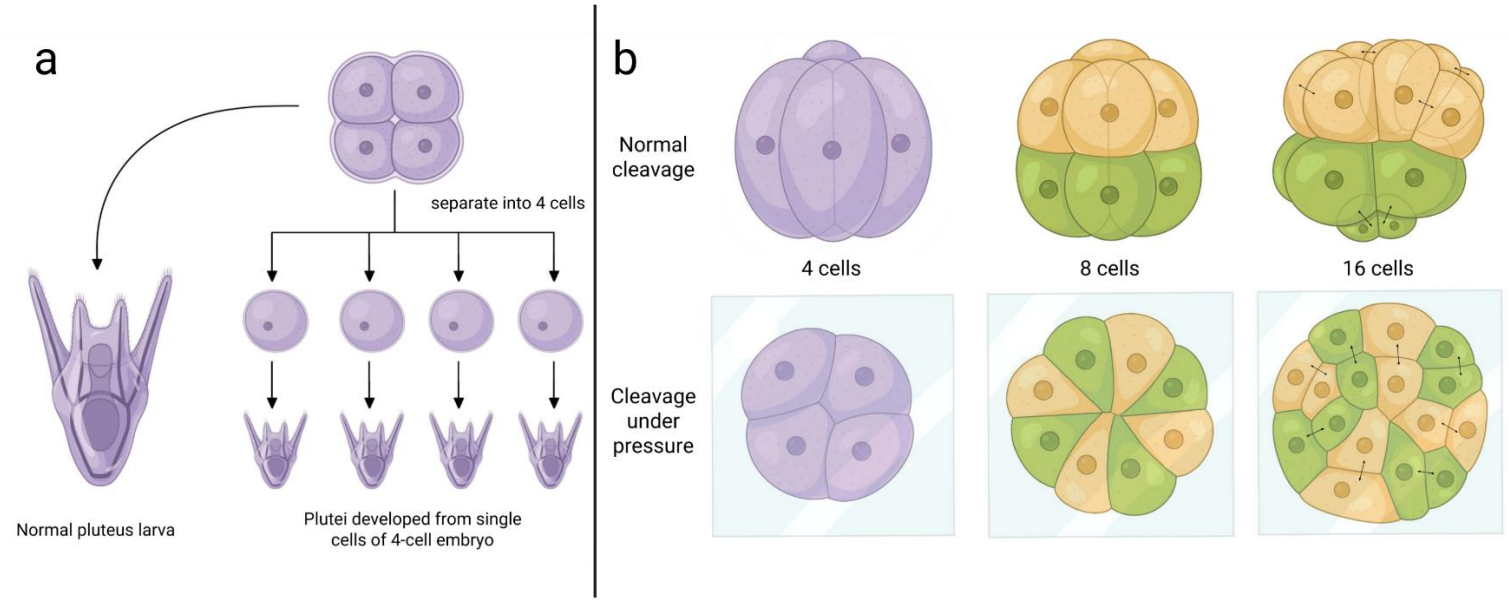

Figure 2.4 | Hans Driesch's experiments. (**a**) A sea-urchin embryo at 4-cell stage can be split into 4 separate cells, each developing into fully patterned yet smaller plutei (larvae). (**b**) (top) Lateral view of normal cleavage for sea urchin embryo from 4 to 8 and 16-cell stages. Due to perpendicular change of division axis, 8 cells are asymmetrically attaining animal (yellow) vs. vegetal (green) pole identities. (bottom) A sea-urchin embryo sandwiched between glass slides (top view) splits into 8 and 16 cells all along the same plane of division under the pressure; resulting in mixture of animal and vegetal pole cells. Arrows at 16-cell stage embryos indicate the cleavage axes.

A notable example of induction is the interaction of the **notochord** and the ectoderm driving the development of the nervous system. The notochord is a component of the mesoderm, but its removal from the embryo disrupts the formation of the nervous system from the ectoderm. This is because the notochord releases signals required by the ectoderm to properly differentiate.

What influences the fate of a cell?

Let us consider here the concept of **'cell type'**. What features are important in determining a cell type – what distinguishes a neuron from a muscle cell? We can list four primary attributes: function, gene expression, shape, and developmental potential. Each cell type in the body plays a specific role. For instance, a neuron transmits signals, while a muscle cell contracts. As such, different cell types differentially expressing sets of genes lead to the production of function and structure with specific proteins. Additionally, different cell types possess distinct morphologies which reflect their function and are a consequence of their expression patterns. Finally, some cells can differentiate into multiple cell types, while others are restricted to a specific lineage.

What ultimately determines cell type are three sequential processes: **induction**, **specification**, and **determination**. Only once the determination is reached, the fate of the cell is committed. Determination is why our liver cells cannot form muscle cells, and why our neurons do not divide into skin cells. Determination locks the fate of the cell and its daughter cells. Cells that are specified but not yet determined still have the capacity to change cell types, although this requires transplantation into a new cellular environment.

However, early differences in cell composition can also arise from intrinsic, not external, factors through **auto-commitment**. The cellular composition of proteins, RNAs, phase separated bodies can be uneven in the egg either prior to fertilization or become so after. As such, the resulting cells from very first divisions harbour different concentrations in their cytoplasm, which in turn can influence gene expression. Auto-commitment is the consequence of polar cells dividing and unevenly distributing their contents.

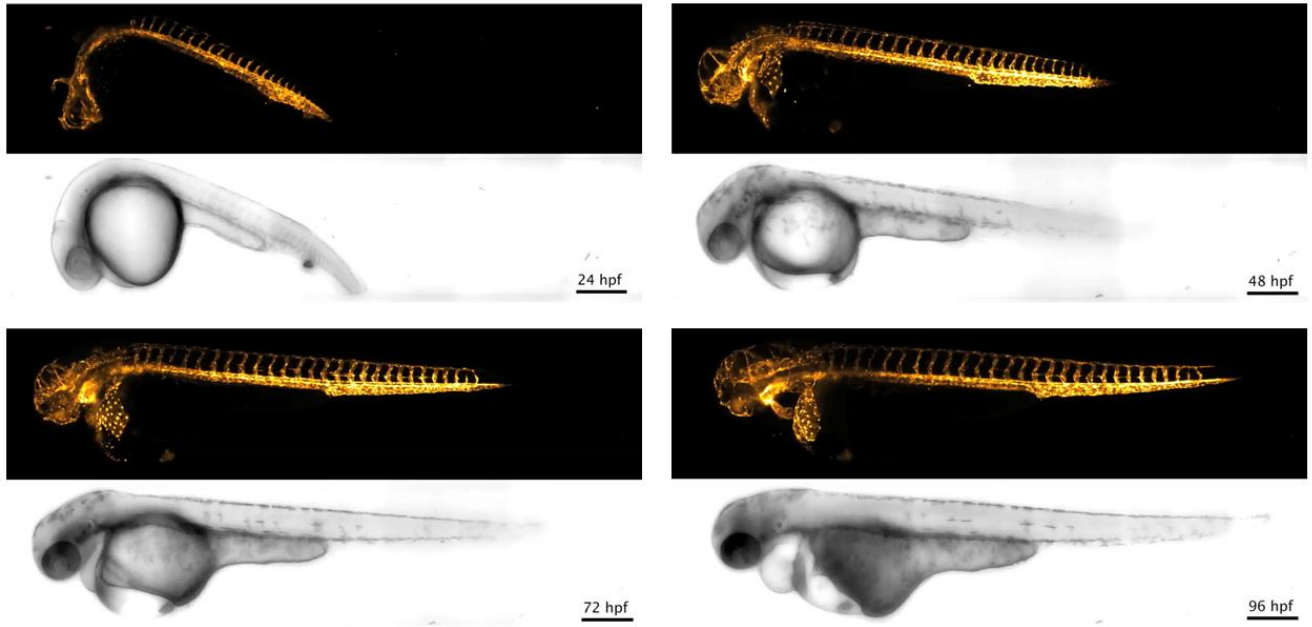

Figure 2.5 | Vascular development of zebrafish embryos over the course of first four days post-fertilization (top-left to bottom-right). Timelapse snapshots are reproduced from light-sheet microscopy in Daetwyler *et al.*, Development, 2019.

*To illustrate the cellular activities driving development, let us focus on the remarkable events unfolding during a timelapse of a developing zebrafish embryo, capturing approximately four days of development (**Figure 2.5**). The marker at the top highlights the vascular cells*

whereas the bottom is transmitted brightfield microscopy. Being observant, let us pay close attention to the vascular cells and consider how this network is formed, as if we were conducting an experiment in a laboratory setting. To set the scene briefly, timelapse showcases the brain vasculature, the blood islands that contribute to the heart's development, and vessels forming along the body axis. The dark areas between the blood vessels represent the somites, which are crucial for the embryo's structure. As the embryo elongates, we see new connections forming.

Now, let us exemplify the four main cell activities relevant to development that were discussed in the last chapter:

1. Multiplication:

Which multiplication activities do you observe in this timelapse? Do you see ***proliferation*** *or* ***cleavage****?*

You are observing proliferation.

2. Cell Movement:

What types of cell movements can you identify?

Migration *is evident as some cells appear to be moving actively.*

Additionally, we can observe ***ingression****, where cells enter and connect with other tissues.*

Motility: *Some cells are highly motile, while others remain stable. This variation in movement is significant.*

3. Induction and Differentiation:

The somites induce surrounding cells to adopt specific fates, such as becoming vascular cells. This process involves ***induction****,* ***commitment****, and ultimately* ***differentiation****. While we cannot see all these processes with the single marker used in this case study, they are indeed occurring.*

4. Signalling:

We cannot directly observe signalling in the timelapse. However, there are two primary signalling mechanisms at play:

- ***Paracrine Signalling:*** *This involves signals diffusing from nearby cells to induce changes in other cells in the vicinity. e.g., fibroblast growth factor (FGF) signalling.*
- ***Juxtacrine Signalling:*** *This is crucial for establishing connections between arteries and veins. In the time-lapse, you may have noticed a structure resembling a pillar forming as arteries and veins connect. This involves leading cells and stalk cells communicating to establish these connections, facilitated by juxtacrine signalling.*

In upcoming chapters, we will delve deeper into these signalling pathways and their roles in development.

FURTHER READING

Chemotaxis in Escherichia coli analysed by three-dimensional tracking. H. Berg, D. Brown, **Nature**, 1972. DOI: 10.1038/239500a0.

The cell in development and inheritance. E.B. Wilson, **MacMillan**, New York, 1911. https://archive.org/details/cellindevelopmen1911wils/.

Reproduction: Antiquity to the present day. Eds. N. Hopwood, R. Flemming, L. Kassell, **Cambridge University Press**, 2018. DOI: 10.1017/9781107705647.

Chapter 3: INDUCTION AND CELL SIGNALLING

"THE EFFECTS OF GRACE DEPEND ON EACH PERSON'S KNACK.

OYSTERS CRAFT PEARL IN APRIL SHOWERS; VENOM IS WHAT VIPERS MAKE."

– I. Beligh Effendi (Ottoman historian and poet, 1668 – 1729).

In the previous chapter, we have begun discussing development as a marvel 'a single fertilized cell' can do using the information contained within itself. The historical path coming to this understanding was not easy and straightforward. We have summarized four main principles for embryonic development as discovered by comparative embryologists of late 19th and early 20th century. The last principle was about how cellular decision-making processes depend on neighbours during development, a term called "induction". We wrapped the chapter by discussing "cell-type" and "fate commitment" concepts. Let us first expand on those terms a bit further here, which will then bring us to the topic for next two chapters: how developmental information is created and interpreted by embryo's cells biochemically, *i.e.,* developmental signalling pathways.

Induction and differentiation

When a single fertilized cell begins to divide, it gives rise to all identical progenies. Those daughters shall distinguish themselves from others along the path leading to a modular and functional organism. Cells end up being different so-called "types" with unique morphological, physiological, and functional feats! How do the cells differentiate?

Cell **differentiation** is a two-tiered process which begins with **induction**. Upon – or more correctly stated in response to – induction, cells commit into being a certain type. **Commitment** involves a two-step process itself: First, a cell specifies into a certain fate, *i.e.,* **fate specification**. Then the cell becomes determined to be that new cell type, **cell determination**. Although we are revolving the idea around cellular differentiation here, we can speak of same tiers for patterning processes. One can say a group of

cells is **induced** to "form" a tissue pattern, which then is carried out as they **commit** to **specify** that pattern and become **determined** to establish the pattern.

The process of differentiation begins with induction. Induction is a collaborative cellular event, where **inducer** cells or tissues release an **inductive signal** to change the cellular behaviour of another cell. This chemical signal is then recognised by receptor on a **responder cell**, which should not only have the chemical-specific receptor but the ability to respond to the signal. This "ability to respond" is known as **competence**. A competent responder cell will commit to specialization and differentiation under a normal healthy developmental pathway. As a side note, the inducer signal can come from the cell/tissue itself, although the more pronounced way we observe is intercellular dependencies.

A classic example for this is how the vertebrate eye is carved from the anterior neuroectoderm. The neural ectoderm as an epithelial sheet of tissue developing under anterior (head) surface ectoderm is induced to be optic vesicle under the influence of neighbouring tissues. Meanwhile, the cells which are **determined** to become optic vesicle would **induce** formation of the eye lens out of surrounding ectoderm next to themselves. If the optic vesicle is dissected out, the lens formation will halt. If a tissue other than optic vesicle gets grafted under the surface ectoderm, it will, as well, not trigger the lens program. Alternatively, if the optic vesicle gets grafted under surface ectoderm of the trunk (not head), the surface ectoderm will be indifferent to any lens induction signal either.

Just as competence for induction depends on cellular position, timing is equally – if not more – crucial. Cleverly designed experiments of "**fate mapping**" enabled classical embryologists to understand this. The go-to fashion of mapping fates is by staining a certain piece (group of cells) of early unshaped embryo with a vital dye and observe how those cells, and their progenies, spread over the course of development. This way embryologists were able to establish fate maps for various model organisms which demarcate boundaries for "presumptive" tissues *e.g.,* presumptive eye, presumptive mesoderm, presumptive heart field, so-on-so-forth. Once the maps are handy, one can further the investigation by transplanting a piece of tissue from an early developing embryo into a different position of a later stage embryo. Why? This way, the embryologist is able to understand whether a piece of tissue "is specified" for a certain cell fate at that given stage or not. For instance, if you dissect the tissue which will eventually form the eye (presumptive eye region) from a gastrula

and engraft it along a late (neurulating) stage host embryo's trunk, you'd observe those engrafted cells now becoming part of somitic tissue instead of eyes. However, if we repeat the same experiment with a later stage (neurulating) host, you would see an eye-like structure growing in the trunk of the host embryo. Such experiments taught us that **development is a gradual process with cells committing into certain fates and differentiate under the influence of inductive signals, which they are competent to respond within a definite time-window.**

Intercellular signalling pathways in development

There are numerous cell-signalling pathways which coordinate the response in cells, alter gene expression and guide commitments. Intriguingly, only a handful of cell signalling pathways are used in developing animal embryos. Why quite a conservative number of pathways evolved for multicellular patterning of animals has important implications from evo-devo[1] perspective. An exhaustive list would not exceed ten pathways, out of which we will focus on five specifically: receptor tyrosine kinase (RTK) signalling, Wnt signalling, transforming growth factor-β (TGF-β) family signalling, Hedgehog (Hh) signalling, and Delta/Notch signalling as pathways responsible for a majority of animal embryo patterning events, from fruit flies to humans. Those can first be grouped into two: pathways sensitive to **paracrine** and **juxtracrine** inductive signals. **Paracrine** signalling molecules diffuse outside of the producing tissue or cell, and can take the form of **lipid moieties, extracellular molecules,** or **non-signalling coreceptors**. These signals are recognized by the receptors of neighbouring responder cells which then activate their cognate cell-signalling pathway. First four pathways listed all belong to that category. On the other hand, **juxtacrine** signalling molecules exclusively relay signals through a cell-to-cell contact dependent manner. Through either homophilic or heterophilic binding of surface exposed glycoproteins, the inducer cell instigates intracellular signalling in the responder cell. The interaction between Delta ligand and Notch receptor in between neighbouring cells is an example.

---

[1] We will expand on this near the end of the book. Briefly, "evo-devo" stands for evolutionary developmental biology. Although from the times of Darwin evolutionary theory has been considerate of embryology, the advent of modern synthesis put the development on the back burner to explain how species evolve. Molecular developmental biology and comparative studies in last decades makes "evo-devo" a flourishing field.

A deep question of development is how so few types of biochemical signals instigate numerous patterning and cell fate commitment decisions for various parts of the embryo and organs. For instance, fibroblast growth factor (FGF) signalling, which is a sub-type of RTK pathway, is instructive in patterning and growth of vertebrate embryo segments, limb buds, heart field, and induction of migratory neural crest cells. The list can easily be extended for this 'jack of all many trades, master of all'. Those are a few snips of developmental patterning taking place in between Carnegie stages 10 and 14 for human embryos within a duration of mere 2 weeks!

Architecture for signal transduction

Regardless of whether a cell is releasing a paracrine or juxtracrine signal, there exists a shared architecture for signal transduction. A **ligand** is presented from a cell and interacts with a **receptor** on another cell. Once the ligand is bound to the receptor, the responder cell produces or activates an **effector molecule** and instigates a **cellular response**. Variations come into play if we exemplify those four layers.

1. **Ligand:** Some ligands are secreted (paracrine) and diffuse into the extracellular space, whereas other act over short distance being membrane-bound (juxtracrine). Some ligands have lipid-moiety *e.g.,* Wnts and Hedgehogs essential for their signalling control. Some ligands can only act as dimers *e.g.,* TGF-β family ligands.
2. **Receptor:** Certain receptors need to dimerize and cross-phosphorylate each other *e.g.,* RTK signal receptors. Some receptors heterodimerize *e.g.,* Type II and Type I receptors of TGF-β family, where former phosphorylates the latter for downstream activity. Certain pathways would also rely on non-signalling receptors *e.g.,* TGF-β family, or signalling receptor pairs *e.g.,* Wnt, for proper transduction.
3. **Effector:** Signalling cascades usually come down to a bottleneck activator molecule which we can call 'the effector molecule'. The levels of the active effector molecule inform the receiving cell a quantitative readout of the external signal. For instance, Akt and Erk are two-branch effectors for FGF signalling, β-Catenin is for Wnt signalling, and Smads for TGF-βs *etc*. Those effectors are activated via post translational modifications *e.g.,* phosphorylation, acetylation, or cleavage. It is essential for cells to sense

the high and low ends of incoming signals properly, regardless of whether they respond to it or not. We call this "signal fidelity".

4. **Response:** The response to external stimuli in receiving cells can also be in different sorts. RTK, TGF-β and Notch signalling result in activation of transcription factors modulating gene expression. However, Wnt and Hh signalling act by removal of repression. Importantly, the final outcome can come in a variety, *e.g.,* post-translational modification of certain proteins or cytoskeletal rearrangements that can change cell shape and motility.

Receptor-tyrosine-kinase signalling pathway

Epithelial, vascular endothelial, neuronal, fibroblast growth factors (EGF, VEGF, NGF, FGF) are some examples for receptor tyrosine kinase (RTK) signalling pathways. Those ligands result in dimerization and subsequent cross-phosphorylation of their receptors at tyrosine residues upon binding. A common theme in RTK signalling is the three-tiered phosphorylation cascade of signal activation. Once the effector molecule, *e.g.,* ERK or MAPK, gets doubly phosphorylated, it can enter the cell nucleus and phosphorylate transcription factors.

Let us detail this for FGF signalling as an example. FGF ligands are trapped by extracellular proteoglycans and presented to the FGF receptor. The receptor dimerizes. Guanine exchange factor (GEF) is activated through membrane localized adaptor protein SOS (Son of sevenless[2]) and then adds phosphate to Ras bound GDP. Ras-GTP can then phosphorylate Raf (Tier 1) which will phosphorylate MEK (Tier 2); and MEK is the kinase of ERK (Tier 3) (**Figure 3.1**).

[2] Those names will keep popping up throughout our developmental biology adventure, mostly pointing to awe-inspiring creativity of *Drosophila* scientists! Expectation is that some of them will make sense by the end of the road.

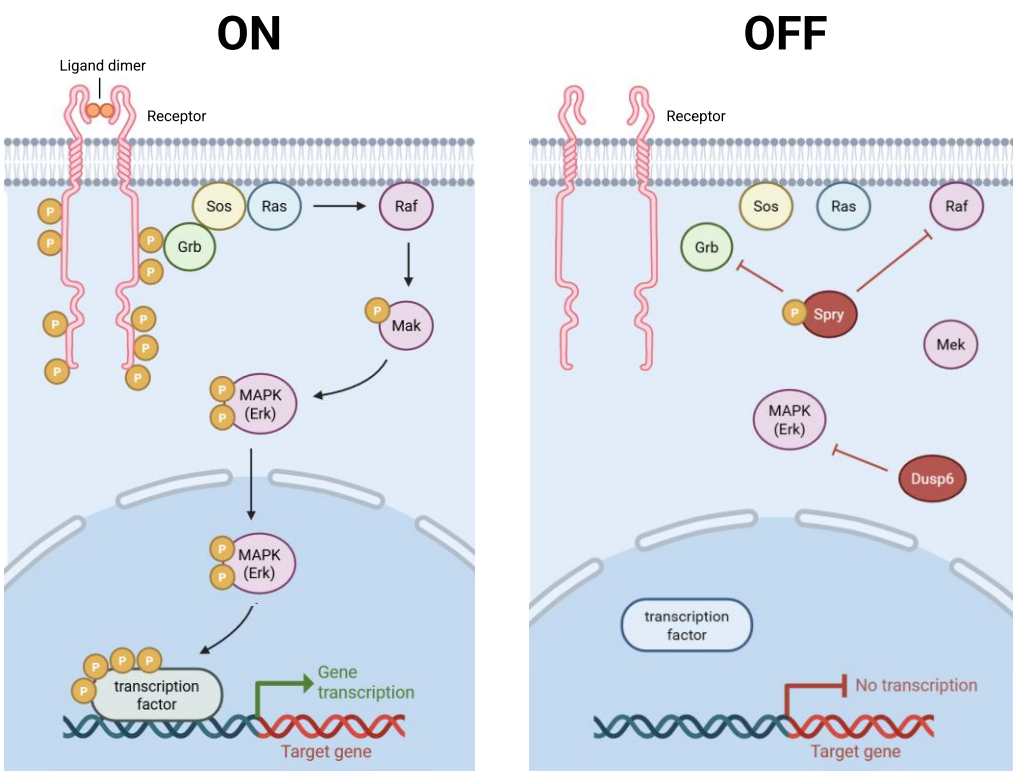

Figure 3.1 | Receptor tyrosine kinase (RTK) signalling pathway.

What is magical with **3-tiers**? Or is there anything magical about it? Indeed, some simple math on differential equations for phosphorylation or dephosphorylation reactions can give us a clue. Let us denote the molecules in these three tiers with X, Y, and Z; Z here being the effector. Let us also add superscript "P" to indicate their phosphorylated state. We can define the reaction rates which convert X to X^P or backwards, Y to Y^P, Y^P to Y^{PP} or Y, Z to Z^P, and Z^P to Z^{PP} and Z (**Figure 3.2a**). Assuming a signal beginning at $t=t_0$ and time evolving the parameters shows us that, as the number of tiers increases, Z can respond to both weak and strong signals robustly with a steep sigmoidal curve (**switch-like response**). Signal activation time can be sped up to a point with this trick, which incidentally happens at three or four tiers. Adding more tiers beyond that becomes superfluous. This lets cellular response to switch on and off for a wide range of external signals, and lack thereof, in a timely fashion (**Figure 3.2b**).

Notably, MAPK (ERK) and MAPKK (MEK) are **doubly phosphorylated** in the cascade. Why so? Again, some math would tell us that if upstream molecule X^P activates a molecule Y, it forms a YX^P complex. Then we can have both Y^P and Y^PX^P complex increasing in expense of the initial YX^P. As Y gets phosphorylated again

leaving X^P aside as Y^{PP}, Y^PX^P complex as well will begin to decline. This double phosphorylation trick will **amplify** even weak incoming signals to a robust amplitude (**Figure 3.2b,c**).

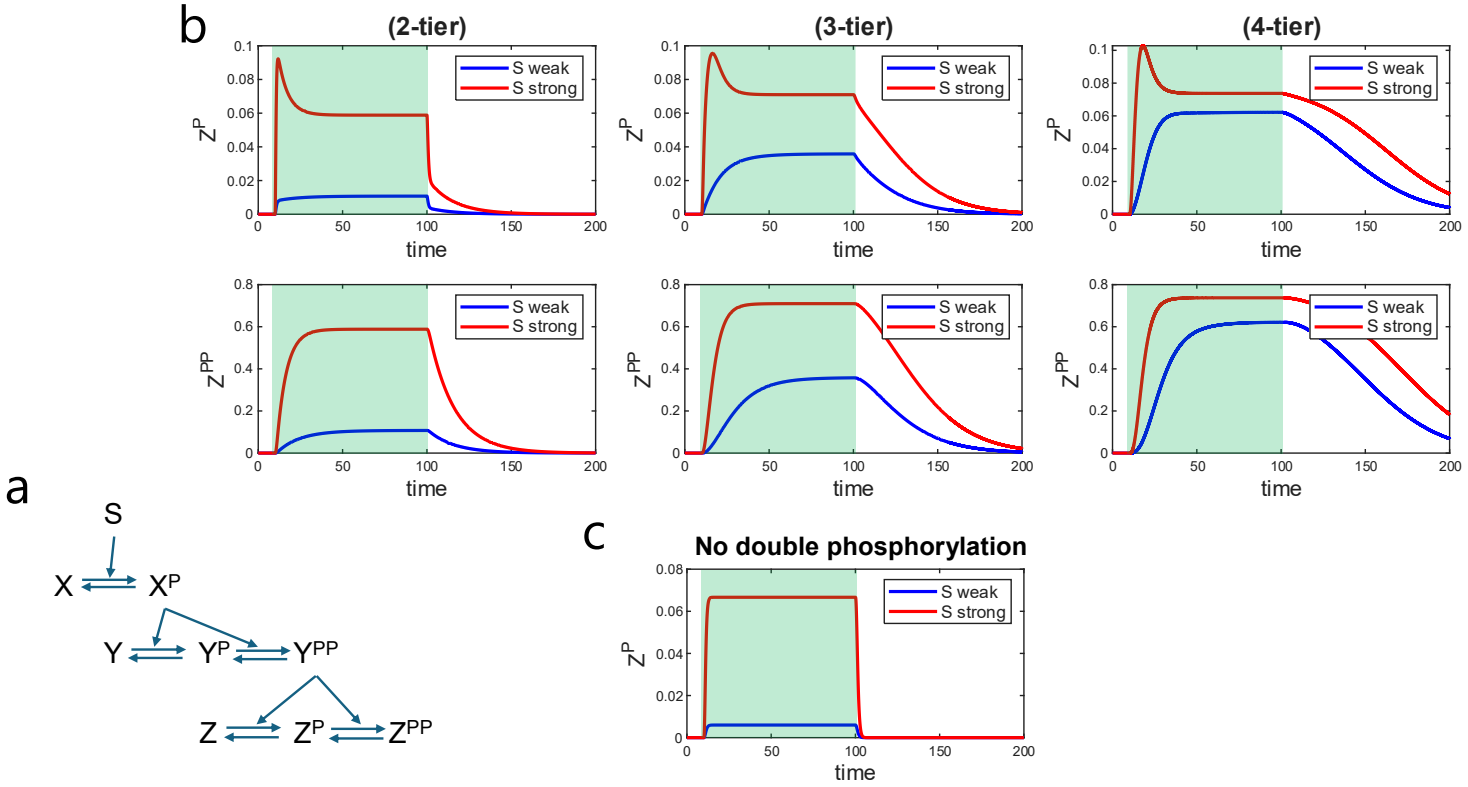

Figure 3.2 | (a) RTK signalling with 3-tiered cascade and double phosphorylation. S is external signal, Z^{PP} is the activated effector molecule as signal readout. (b) 3-tiered signal cascade (middle) enables robust ON (green shade)/OFF switch-like response to the incoming signal, while still being able to distinguish weak (blue, S=5%) from strong (red, 100%) signal. (b) In the absence of double phosphorylation of the kinases, signal amplification is drastically weak.

Wnt signalling pathway

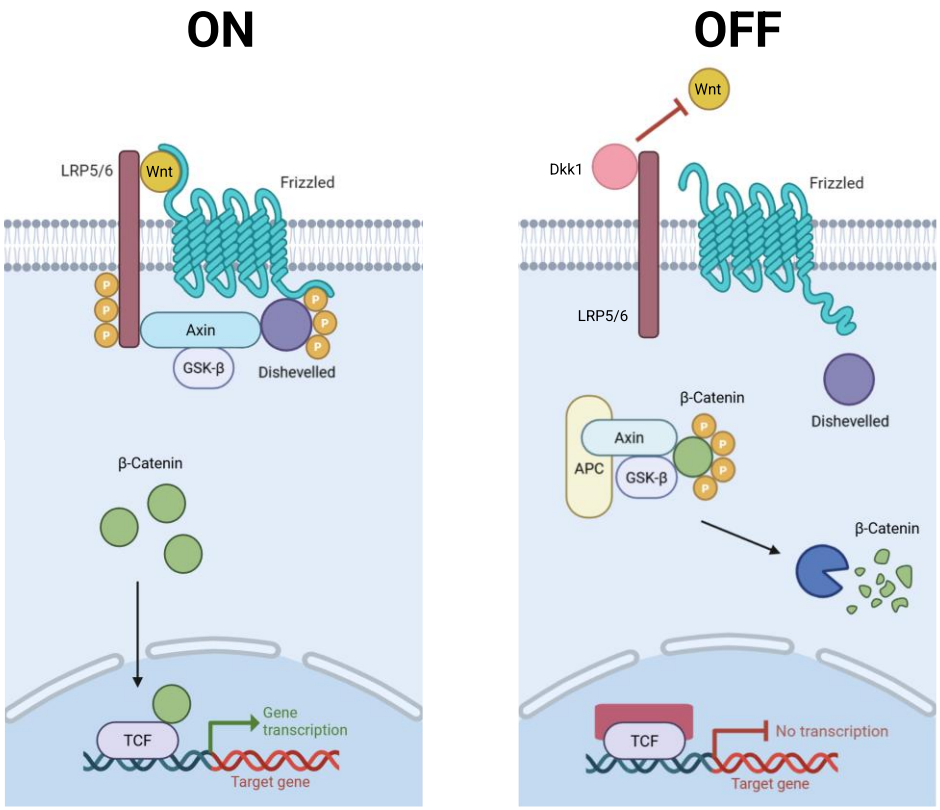

Figure 3.3 | Wnt signalling pathway.

Wnt (pronounced as "wint") gets its name from a merger between two genes: *Int*, which was discovered in mice, and *wingless*, which was discovered in fruit flies, *Drosophila melanogaster*. Scientists later figured genes coding for Int and Wingless, first being central in cancer progression and the second causing wing loss if mutated in flies[3], are homologous.

Frizzled and LRP are the co-receptors for Wnt ligands (**Figure 3.3**). Once ligand-bound, phosphorylated Frizzled can keep the Axin-GSK-β-CK1γ destruction complex at the membrane, preventing the phosphorylation mediated degradation of Wnt signal effector molecule β-Catenin. In the absence of external stimuli (no Wnt ligands), Wnt effector β-Catenin is produced but is continuously degraded by the destruction complex, preventing its entry into nucleus. In active signalling, β-Catenin rescued from the destruction can enter the nucleus and act as a transcription factor together with co-factor TCF. Similar

[3] It is a common practice to name a gene with the phenotype its loss results in. That doesn't necessarily mean the *wingless* gene is the "cause" of wing development. However, losing *wingless* is one of the ways of losing the wing development in *Drosophila*.

to two features mentioned for RTK signalling, there are distinct tricks making the signalling architecture robust in Wnt and other signalling pathways. A central feature for all pathways is the feedback regulation. Let us divest from our discussion of developmental signalling pathways and dive deeper into the 'feedback control' a bit as a case study.

Case Study: Feedback Regulation of Developmental Signalling

When considering feedback control mechanisms, we are usually introduced to an overly simplistic sketch where an initial transcription factor activates the expression of a hypothetic gene one, which itself encodes for a secondary transcription factor to activate gene two. The product of gene two transcription and translation is a tertiary transcription factor that further activates genes one. Through this circular cascade, the response from that initial activator is reinforced. This ***positive feedback*** *ensures that the cell can respond to even a weak primary signal. Likewise, we can consider the tertiary transcription factor out-competing the first activator in binding to the regulatory domain of gene one. In that case, a* ***negative feedback*** *loop will establish. As a positive feedback system can reinforce a weak signal, a negative feedback loop can shut off responses or stabilize them when a signal persists. Why is this important biologically? Consider if there is a high background rate of any given "signal", the cell may not want to consistently respond to it and a negative feedback system will act to dampen the response. However, it is essential to keep in mind that picture is quite simplistic. Both positive and negative feedback can take place through gene activation or through post-translational modifications that alter protein function or state. We will go through examples of negative and positive feedback in the important signalling pathways during embryonic development.*

Negative feedback in FGF signalling

Fibroblast growth factor (FGF) is a type of RTK signalling pathway. Cascade begins with the FGF ligand binding to the membrane bound FGF receptor (FGFR), which activates a larger multiprotein complex. This complex includes the ligand-bound FGFR, Son of Sevenless (SOS), Frs2 and Grb2. Collectively, this complex activates Ras which in turn phosphorylates Raf; the latter then phosphorylates MEK whose target is the extracellular signal-regulated kinase (ERK). Once doubly phosphorylated by MEK, the ppERK translocates into the nucleus and phosphorylates a variety of transcription factors. These activated transcription factors drive the expression of many genes. FGF signalling has been shown to turn on hundreds of genes, and a good number of them in turn act to inhibit the FGF signalling cascade. We will briefly exemplify four gene products: Dusp1, Dusp6, Sprouty, and Sef. As the FGF signalling activates the expression of these proteins, each of them targets unique components of the pathway to inhibit the propagation of the FGF signal. Dusp1 functions to dephosphorylate ppERK into ERK within the nucleus; Dusp6 dephosphorylates ppERK outside of the nucleus. On the other hand, Sprouty and Sef hit upstream the cascade. Sprouty prevents action of Raf, as well as formation of Frs2 and Grb2/SOS complex. Sef both cuts off the extracellular signal from the receptor and prevents MEK's kinase action on ERK. ***Overall, post-translational and transcriptional negative feedback loops ensure timely inactivation of the FGF pathway.***

What happens when the ON/OFF switch of FGF signalling does not function timely? A drastic example is the mutation in extracellular domain of FGF Receptor 3 (FGFR3) in human patients, resulting in premature activation of the receptor and is associated with congenital defects. Mutated FGFR3 prematurely triggers chondrocyte fate commitment, the cells which

*give rise to the cartilage tissue, resulting in lower counts of them in the end. In result, cartilage development stops before the full growth of the fetus until birth. Failing to give rise to enough cartilage for bone development, the patients carrying the mutation are unviable after birth, as their ribcages are often too small to support breathing (**thanatophoric dysplasia**).*

Negative feedback in Wnt signalling

The Wnt signalling cascade begins with the Wnt ligand binding to the Frizzled (Fz) receptor which sits in a multiprotein complex with the Lgr and LRP co-receptors. This ligand bound Lgr-Fz-LRP complex functions to keep a larger protein complex, which includes proteins Axin, APC, GSK3β bound together at the membrane level. Additionally, ligand bound Lgr-Fz-LRP recruits the Dishevelled protein (Dvl in vertebrates / Dsh in invertebrates) to this complex. Dvl plays an important role and blocks the activity of GSK3β. What is the activity of GSK3β in the absence of Wnt ligand binding and Dvl recruitment? GSK3β is a kinase that targets β-Catenin, a transcription factor that is degraded by the cell when phosphorylated. Therefore, when the Wnt signal binds to Fz and Dvl is recruited to this large protein complex, β-Catenin will remain unphosphorylated, translocate into the nucleus, and activate the expression of many genes, including axin2, rnf43, and ddk1. Similar to Fgf signalling, gene products are inhibitors to the Wnt signalling cascade. Axin2 acts both at a cytoplasmic and membrane level to inhibit Dvl from interacting with GSK3β; Rnf43 localizes to the membrane and drives Dvl degradation; Dkk1 is active extracellularly and interferes with the Wnt-ligand binding to Fz.

TGF-β signalling pathway

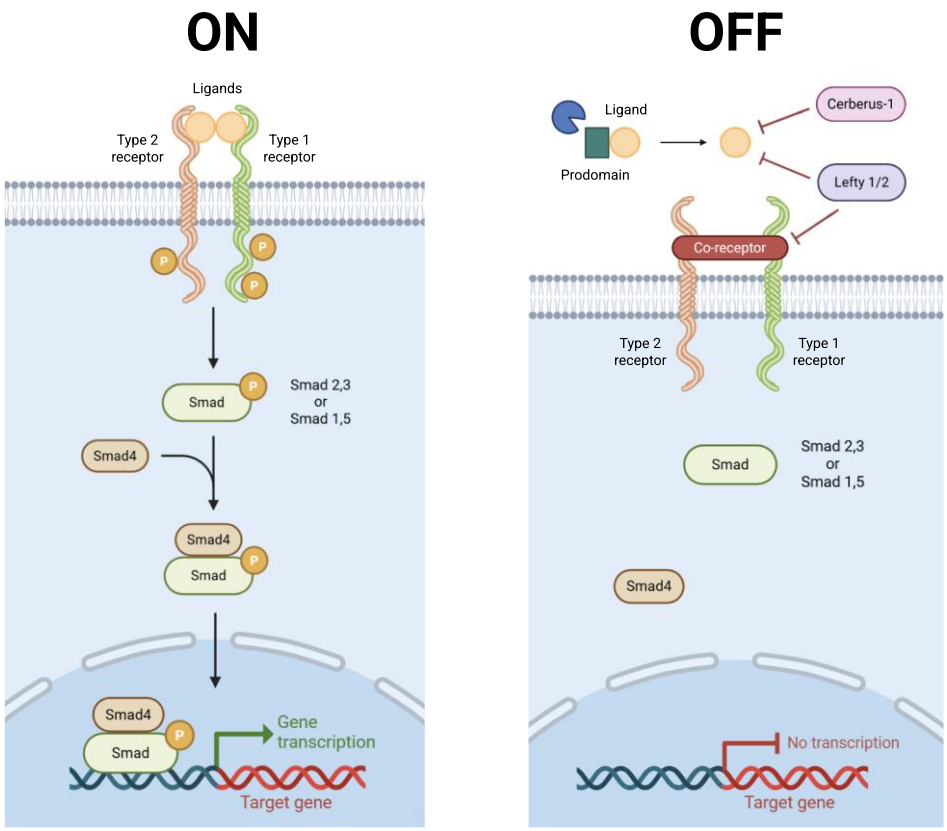

Figure 3.4 | Transforming growth factor (TGF)-family signalling pathway.

We had previously brought up that the ligand of TGF-β acts as a homodimer which can bind to two different types of receptors: Type I and Type II. Following dimerized ligand binding, receptor Type II will become phosphorylated, which in turn results in the phosphorylation of receptor Type I (**Figure 3.4**). Subsequently, two cytoplasmic proteins, Smad2 and Smad3 (Smad2/3) become phosphorylated which allows them to interact with Smad4; this interaction enables the Smad complex to enter the nucleus and directly activate the expression of target genes as a transcription factor. If we compare and contrast, this is similar with what we see in Wnt signalling, where the effector molecule localizes to the nucleus acts as a transcription factor. Consider the alternative, which we discussed for RTK signalling, where the ppERK effector molecule is moved to the nucleus to activate other transcription factors, rather than being the activator molecule itself.

The TGF-β signalling cascade is very diverse at the ligand and receptor level, owing to the TGF-β proteins belonging to a very big protein super-family. Notably, TGF family is associated with a vast variety of

developmental processes, including bone formation, left-right asymmetry, and axis establishment. How is that a "single" signalling pathway can control such diverse processes? Well, this is rooted in the diversity of TGF-β ligands, and the different Smad proteins they activate. Smad2/3 are activated by the ligands Activin, Nodal, and TGF-β, but Smad1/5 are activated by another TGF-β family protein, the bone morphogenic protein (BMP). This **distinction in effector molecules** is important to understand how different ligands may have different downstream consequences in terms of embryonic development.

I hope the discussions so far on three key signalling pathways may convince us to call them both as a jack of all trades (contributing to many distinct developmental processes) and a master of all (these pathways are very tightly controlled, regulated, and reproducible during embryogenesis). To establish the robust patterning in embryogenesis, one requires very tight control and regulation of a given pathway. A cell should respond to a signal at the appropriate location and proper developmental time point and should be able to overcome the "noise" flowing from their extracellular environment. This requires a multi-layered control of a given cascade – if this were not the case, we would expect to see birth defects or developmental problems crop up at a far greater frequency than we see them observed. We will use the Nodal TGF-ligand activation as an example for the tight layers of control that are observed in the TGF signalling pathway.

Case Study: Fine-tuning in Nodal Signalling

There are many layers of control to ensure that Nodal signalling is carried out in a tightly regulated manner. The first layer of control observed is in the extracellular space, prior to receptor binding. When the Nodal inducer signal is secreted, the protein possesses a ***prodomain*** *deeming the ligand unable to bind its receptor. Only cells with the capacity to proteolytically cleave Nodal and expose the functional ligand domain can respond to the signal. Additionally, the Nodal signalling also encodes inhibitors to the pathway, such as Cerberus-1 and Lefty1/2, establishing* ***feedback regulation****. Cerberus and Lefty are secreted to the extracellular space and can prevent ligand-receptor binding. The prodomain of Nodal*

and the secretion of inhibitor proteins act as forms of regulatory control to Nodal which takes place even before the signal enters the cell.

These "control" points are important to maintain a tight regulation – which is especially important in TGF signalling since the cascade is more direct than the other examples we have gone through. Following the phosphorylation of the Smad proteins, a cell will activate gene expression. Unlike Fgf signalling, where we see a tiered cascade response, activation by TGF ligands is relatively straightforward. However, there is some additional degree of control once the signal is inside the cell. Once the Nodal signal has entered the cell, and the transcription factors have been activated by phosphorylation and transported to the nucleus, the cell turn on the production of more Nodal ligands. This creates a ***positive feedback*** *system which allows for signal reinforcement. If certain Nodal signalling centre is established in an embryo it can fortify the signal, and more cells in that vicinity can turn on the Nodal signal. This is important in the asymmetric patterning of the embryo; ensuring that the heart is one on side and its loops are properly developed while the liver develops correctly on the other side of the axis. We will delve deeper into those mechanisms in upcoming chapters. Notably, there are post-transcriptional layers of control for nodal transcripts which rely on* ***microRNA****s. To sum, TGF signalling cascade here uses a combination of negative and positive feedback systems to regulate both the individual cells response to an extracellular signal and the development of tissues and the embryo.*

An important and reoccurring theme in developmental biology is the central role of signal inhibitors. **Surprisingly, many of the molecules that play key roles in shaping the vertebrate body plan act as extracellular antagonists, or inhibitors, of intercellular signalling molecules which block and modulate**

the signalling levels. Some examples are Cerberus, Cerberus-like (Cbl), and Lefty for Nodal; Noggin and Chordin for BMP; and Dickopff and Frizzled-B (Frzb) for Wnt signalling pathways.

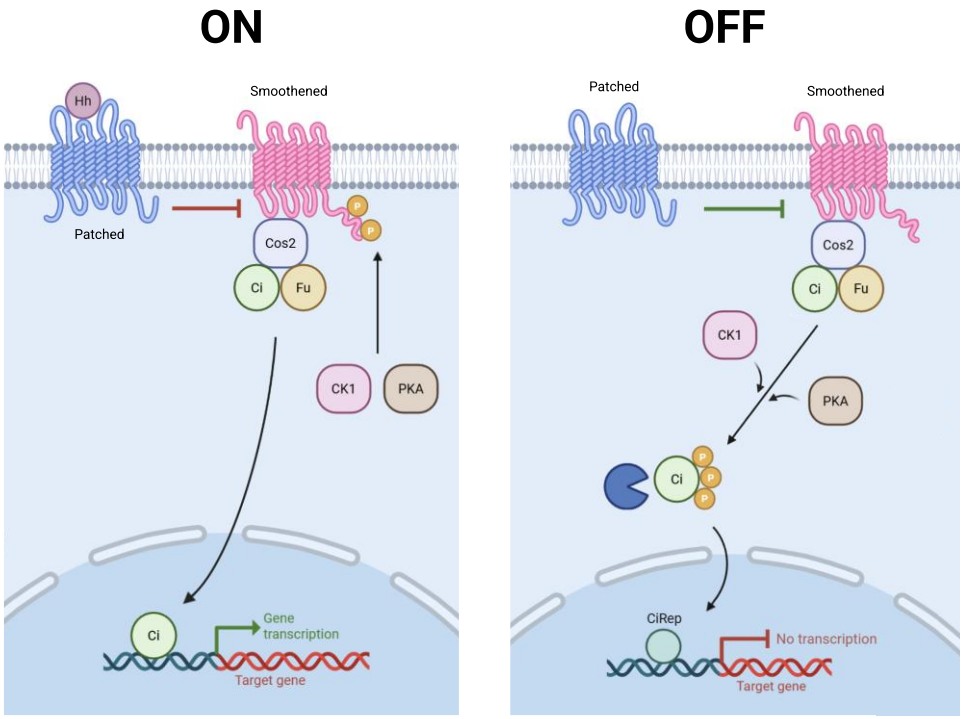

Figure 3.5 | Hedgehog signalling pathway.

The Hedgehog/Shh signalling pathway

The Sonic hedgehog (Shh) signalling pathway was identified through a fruit fly genetic screen for regulatory genes of development. Scientists who carried out those screens in early 1980s, Christiane Nüsslein-Volhard and Eric Wieschaus, identified numerous essential genes for *Drosophila* development including *hedgehog* and were awarded the Nobel prize in 1995. The *hedgehog* gene was first identified because it disrupted the repetitive cuticle segments on the fruit fly larvae, making them look like 'hedgehog's with denticles covering whole larvae across cuticles. We now know it activates a signalling cascade that is important for numerous tasks during embryonic development. Three different vertebrate homologues are found for the Drosophila *hedgehog*: *Sonic hedgehog* (*Shh*), *Indian hedgehog* (*Ihh*), *Desert hedgehog* (*Dhh*); first one with namesake from the famous video game character popular in early 1990s when the homologue was discovered.

Broadly, the hedgehog (Hh) signalling molecule binds to the transmembrane receptor Patched (**Figure 3.5**). In its apo state, Patched inhibits the membrane localized Smoothened protein. Smoothened maintains a multiprotein complex at membrane level which keeps the functional effector molecule of Shh signalling, Gli (vertebrates) and Cubitus interruptus (Ci; fruit flies), in an interaction with the microtubules of the cell. Ci/Gli can only leave this site following phosphorylation and proteolytic cleavage. It is only when Patched is bound by its ligand that Smoothened is functional and activated. In the Smoothened activated state, Gli/Ci bound to the microtubules can leave freely to the cytoplasm. Both the full length and cleaved Gli/Ci can translocate to the nucleus, but their state will determine whether the transcription factor will act as a repressor (cleaved) or activator (full-length) to the same set of genes.

1. Lipid modifications of signalling ligands

An important feature of the pathway is the **post-translational modification** of the ligand. When Shh is translated, it is post-translationally modified by addition of a lipid-moiety groups, cholesterol and palmitic acid. Depending through which apparatus Shh is secreted, it may be a monomer, a multimer, or be released in a lipid exovesicle. Lipid modification of signalling ligands is not unique only to the Shh pathway; in Wnt, we also see that the ligand is modified through palmitoylation. Why is this happening, biologically? We have only recently understood that the lipid modifications are extremely important to control the extent of cell signalling. Firstly, modification with a lipid helps control the site of secretion; as opposed to the ligand being translated and secreted from a random spot, it can instead be secreted in a group of exovesicles through a dedicated apparatus. Furthermore, we know that the diffusion ranges are different depending on the state of the ligand (monomer vs. exovesicle). A monomer, for example, will have a shorter diffusion range because the cholesterol modification will make it "stick" immediately to the neighbouring membranes that it may encounter. Therefore, if a cell secretes a greater proportion of monomer Shh, this forces the signal of the inducer to be confined to a shorter distance. However, the exovesicles can carry the Shh signal to a greater distance stably. Altogether, lipid modification and variable forms of ligand secretion allow for a cell to "tune" the **diffusion range** and intensity of the signal it is secreting.

Altogether, lipid modifications result in a tighter control of the cell signalling gradients, *i.e.,* how the gradient goes down in the tissue from a signalling centre. We can exemplify this for Shh signalling across

several tissues it patterns in a well-restricted fashion. An example is patterning of the dorso-ventral axis of neural tube specifying certain neuronal fates. For instance, when our central nervous system is forming, we should have our motor neurons only differentiating in a restricted space from the ventral side of the axis, to connect to the muscle. And that is controlled by Shh. Another example, in the wing disc of fruit flies, the top and bottom of the wing structures are split by the restricted Shh signalling centre. A third example we can mention is sub-specification of somites, repetitive block of vertebrate body plan, into certain fates in Shh-dependent fashion. One interesting thing we know, is that some patterning processes, like the mouse limb bud, do not deploy lipid modifications of Shh to increase the diffusion range of the signalling molecule. We do not quite understand how one tissue does it this way and another differently; but we know that lipid modifications impact patterning of signal gradients and tissue development.

2. Early discovery of Hedgehog signal inhibitors and teratology

Early clues for the sonic hedgehog signalling pathway were discovered through an interesting story. In 1950s, ranch keepers in Westerns plains of the United States recognized recurring developmental abnormalities in their livestock. Ewes were born with cyclopia, a congenital disorder where the facial midline does not pattern properly leaving most of the offspring with a single central eye. **Cyclopia**, extreme holoprosencephaly, has long been reported in books related to human developmental abnormalities since Middle Ages. Researchers surveyed and carried out feed trials over the course of seven years what could have been causing those abnormalities in ewes; eventually identifying that the corn lily, *Veratrum californicum*, which was in the diet of grazing animals, produced a compound named as “cyclopamine”. Cyclopamine was a potent inhibitor of the Smoothened receptor. By interfering with Shh signalling, cyclopamine had resulted in sheep embryos with aberrant development of facial structures and limb deformities. This was also one of the pioneering works in teratology which deals with environmental and chemical factors leading to awry embryonic development.

Case Study: Synthetic Developmental Biology

Scientists learning about the regulatory architecture of signalling can now emulate those strategies in nature, to perform the developmental tasks synthetically. A beautiful example is recently published in the journal Science, by Starpornwongkul and colleagues in J. P. Vincent's lab. BMP homolog, decapentaplegic (Dpp) is a central pathway for fruit fly development. In the Dpp mutant fruit fly, fifteen organs including wings are awkwardly crumbled and atrophied, lacking structure: hence the name "deca-penta-plegic". BMP was later identified being homologous to Dpp, where in vertebrates this protein family was known to play a role in bone development: hence the name "bone morphogenic protein". Using the fruit fly as a model for BMP signalling in development, researchers aimed to parse the importance of the initial ligand signal in patterning and growth of fruit fly wing disc with a synthetic biology perspective. They facilitated small antibodies, known as nanobodies, as receptors to the green fluorescent protein (GFP) and designed a synthetic BMP-cascade using GFP as ligands. Expressing chimeric versions of Receptor Type I (Thickveins, Tkv) - GFP nanobody and Receptor Type II (Punt) - GFP nanobody, the experimentalists found that extracellular GFP was able to restore proper organ development in Dpp mutant fruit flies. These experiments revealed that ***diffusion*** *of ligands is important for regulation, as GFP was synthetically expressed only at a local source, yet nevertheless the wing disc patterned and developed properly. Furthermore, the study revealed the* ***non-signalling receptors*** *being essential to keep the diffusive ligands within the developing organ system in 2-D, which are expressed in proximity to and maintain ligand availability for the functional receptors.*

Notch Signalling

Our last station in key developmental signalling pathways is Notch signalling. This also is distinctly a juxtacrine (cell-to-cell contact dependent) signalling pathway, unlike the previously discussed four paracrine (extracellularly secreted and diffusive) signalling pathways. A brief historical background is helpful. Notch signalling was first found in *Drosophila* screenings in the last decade of 20th century. The phenotype was first noticed looking at the patterning of fruit fly wings. In wild-type fruit flies, the edges of their wings harbor specialized cells which are placed with a certain periodicity around the wing. This patterning is important, as these cells play integral sensory roles. Some are mechano-sensory, allowing the fly to respond to the wind and other mechanical forces, while others are chemosensory. These latter cells work similarly to a human nose. However, mutant flies heterozygous for *notch* allele exhibit a disruption to the patterning of these sensory cells and additionally the periphery of the wings display notched-shape defects. The gene that mediates the proper patterning is referred to as *notch*, a namesake for the original mutant phenotype.

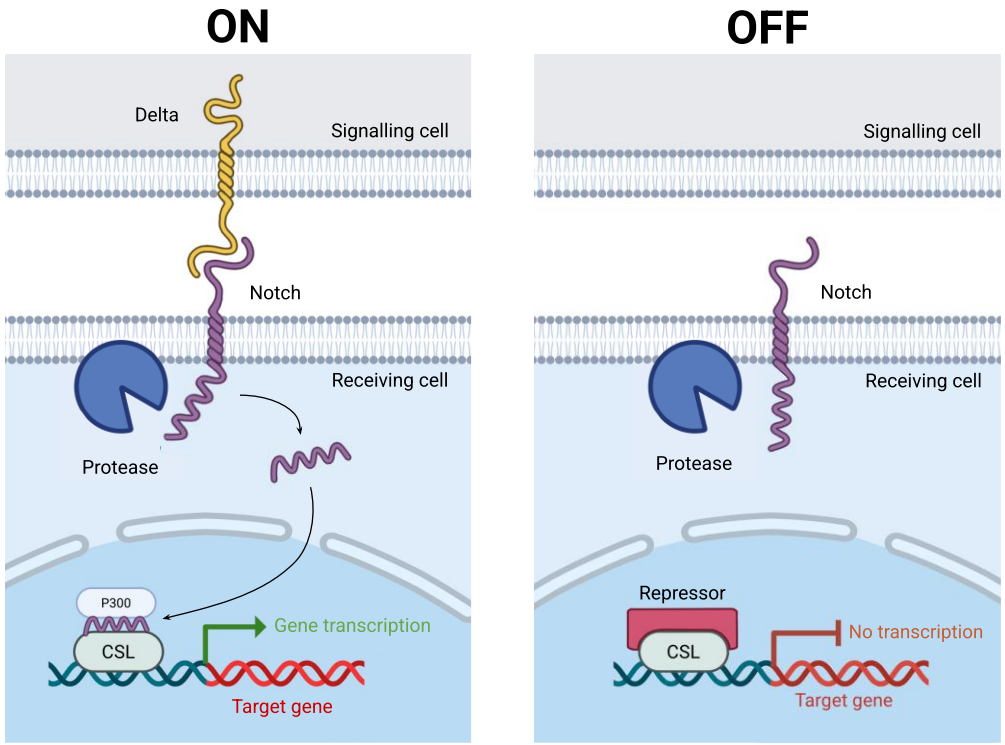

Figure 3.6 | Notch signalling pathway.

The allele *notch* encodes a multi-domain transmembrane protein which is the Notch receptor. This receptor acts in a juxtracrine signalling pathway that helps impart periodic patterning and structuring of organs during development; in fruit flies we see it is important for the spacing of sensory cells on the wing, but Notch signalling is just as important in other animals, including humans.

As a brief reminder, **juxtracrine signalling** is facilitated by the recognition of two membrane-bound proteins that can form either heteromeric or homomeric complexes. As such, this signalling pathway requires cell-to-cell contact for activation. The Notch receptor recognizes the Delta ligand of a neighbouring cell (**Figure 3.6**). Once the neighbouring cell's Delta binds to Notch, the physical force of this interaction results in a conformational change to the cytoplasmic, or soluble, domain of Notch in the signal-receiving cell. This acts as the first "trigger" for the pathway. The next step of the pathway is mediated by a cytoplasmic-facing membrane bound protease complex known as γ-secretase. As Notch undergoes conformational changes mediated by Delta-Notch interactions, the γ-secretase is then able to proteolytically cleave the soluble domain, releasing it from the membrane-bound Notch. The liberated Notch intracellular domain (NICD) translocates to the nucleus where it displaces a transcriptional repressor complex. In coordination with several other proteins, including CSL (RBPJ-κ, in mammals, Suppressor of hairless, Su(H), in fruit fly) and Mastermind, NICD forms a transcriptional activator complex for a variety of genes.

Comparing Notch to the other signalling pathways – like the three-tiered cascades mentioned for RTK signalling – the structure of the signalling cascade is too weak to robustly transfer information. A single molecule gets cleaved off a membrane-embedded protein, and then directly goes to the nucleus. How does the cell control the output of this signalling pathway? A recently proposed model is **"the bind and discard model"** for NICD, such that when NICD binds to the activation site in the nucleus, it would very quickly be tagged for rapid degradation. This leaves a single molecule doing a single task and is a nice measure of how much Notch signalling comes from outside and how much a cell responds to that signal. Once NICD acts as a transcriptional activator, cells no longer need to keep that NICD molecule around.

1. Lateral inhibition mechanism

An additional layer of control is mediated by the genes that are activated by NICD. A major family of proteins activated by Notch signalling is the *hairy-enhancer-split related* genes (*her/Hes*). The Her/Hes proteins function as transcriptional repressors; main targets of which are their own genes as well as *delta*, the gene coding for the Notch ligand, and other Notch ligands *Delta-like* (*Dll*) and *Jagged* in mammals. This means that for cells where Notch signalling is activated, there will eventually be less production of the Delta ligand. And so, Notch activated cells restrict the activation of their neighbouring cells in a tug-and-war process known as **lateral inhibition**. This ultimately results in the neighbours of Notch-activated cells exhibiting a different transcriptional profile and committing to a different cellular fate.

How does a cell "decide" whether it expresses Notch or Delta in this tug-of-war? If you think about the nature of gene expression and transcription, it is a stochastic process referred to as transcriptional noise. When a gene is turned on, either through the removal of a repressor or the presence of an activator, the cell is able to make RNA copies of that gene, which means that it can burst "on" to make these copies and then burst "off" to no longer generate them. These OFF and ON bursts generate a stochastic pattern of gene expression. We can consider this concept of transcription in the context of Delta-Notch signalling. If we have two neighbouring cells, it is possible that both will turn on Delta expression in a stochastic fashion, and that means they will also impart Notch signalling to the neighbouring cells. And, through lateral inhibition, what happens is that as one cells' Notch signalling increases, the surrounding cells will have lower activation levels as explained. At a developmental level, the cells with high Notch signalling will commit to a different fate than their neighbours, explaining how Notch drives the alternating patterning first observed in the sensory structures surrounding the fruit fly wing.

Sparse and alternating labeling for cell fates within a tissue is an essential for development. Why is lateral inhibition a good strategy? If you want to make an organ with repetitive structures, thinking back to the sensory organs of the fruit fly wing, you do not want all your cells in the neuronal ectoderm committing to the same fate. Spaced patterning is managed by the stochastic nature of transcription and the Notch signalling pathway.

2. Notch signalling across development

As previously mentioned, Notch signalling orchestrates more than just the patterning of fruit fly sensory cells. For example, in vertebrate eyes, Notch plays a central role in retina development. The retina is patterned with the neural epithelium cells committing to different neuronal fates. Notch drives the specification of the Müller glia (MG), which is responsible for the physical strength of the retina. These glial cells are situated between the photoreceptors and helps balance the tensile stress as the organ grows. When observing the early retinal development of the zebrafish, researchers tracked Notch signalling through a fluorescently tagged indicator protein, Tp1-Venus. As the structure begins to develop, many of the apical neuronal epithelial cells are observed to be high in the Notch signal. But, as development goes on, the MG cells become interspersed between developed retinal cells. Lateral inhibition through Notch signalling mediates this patterning.

Notch signalling gets slightly more complicated when we look at vertebrates, especially mammals. Mammals have multiple copies of Notch receptors (Notch 1 – 4), and different ligands including Delta-like (Dll 1 – 4) and Jagged (Jag 1 – 2). These diverse ligand receptors control distinct developmental processes; for instance, Jagged in vertebrates is important for blood vessel formation, and Dll proteins in mammals are more important for neuronal fate commitments. In all cases, the NICD is the major signalling molecule that instigates changes in the transcriptional profile of a cell.

In humans, Notch signalling plays a role in the development of many patterned organs, including the kidney, heart development, pancreas patterning, *etc*. Like with fruit flies, the development of some of our sensory organs arises from Notch signalling, specifically, the sensory cells in our inner ear. Notch signalling is also incredibly important for angiogenesis (blood vessel connections), specifically when the arteries and veins must connect with each other. As the blood vessels are connecting, if you want to properly form a connection, you will need some cells to leave the vein and a cylindrical patch of cells to stalk the leading tip cell in this collective migration. This distinction between the leading cell and the stalk cell roles is coordinated by Notch signalling.

3. Non-localized role of Notch in repetitive tissue patterning

Does Notch signalling only exert a localized effect on cell-fate commitment? Although in many cases Notch works through lateral inhibition, it plays a secondary role that is slightly more complicated and requires three criteria to be met following transcriptional activation:

1. The mRNA produced has a short half-life.
2. The protein translated from the mRNA also has a short half-life.
3. The protein translated shuts off its own transcription (negative-feedback loop) with a time delay.

If these conditions are not met, Notch signalling will only induce lateral inhibition in neighbouring cells. However, when all these conditions are met, Notch allows for sustained **intracellular oscillation** of the short-lived RNA and protein levels. This is what we observe for the Her/Hes transcriptional factors that are expressed even in the absence of Notch signalling activation. We see that both protein and RNA levels go up and then drop down within a cell with established time-delayed negative feedback loop. But Notch signalling coupled with Her/Hes expression causes the intracellular "signal" oscillations to enter the same phase and **synchronize** in between many cells. This allows for information to travel in the developing tissue through the generation of a kinematic wave, akin to stadium waves.

This wraps up our discussion of key developmental signalling pathways. Readers may now revisit the "Architecture for signal transduction" subsection on page 32 as a quick summary of the topic.

FURTHER READING

1998 Warkany lecture: Signalling pathways in development. J. Gerhart, **Teratology**, 1999. DOI: 10.1002/(SICI)1096-9926(199910)60:4<226::AID-TERA7>3.0.CO;2-W.

The mammalian MAPK/ERK pathway exhibits properties of a negative feedback amplifier. O. E. Sturm *et al.*, **Science Signalling**, 2010. DOI: 10.1126/scisignal.2001212.

Wnt/β-catenin signalling, disease, and emerging therapeutic modalities. R. Nusse, H. Clevers, **Cell**, 2017. DOI: 10.1016/j.cell.2017.05.016.

Patterning and growth control in vivo by an engineered GFP gradient. K. S. Stapornwongkul *et al.*, **Science**, 2020. DOI: 10.1126/science.abb820.

Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain. E. Schroeter *et al.*, **Nature**, 1998. DOI: 10.1038/30756.

Pattern formation by lateral inhibition with feedback: a mathematical model of Delta-Notch intercellular signalling. J. R. Collier *et al.*, **Journal of Theoretical Biology**, 1996. DOI: 10.1006/jtbi.1996.0233.

Chapter 4: FERTILIZATION

"WE MAY HAVE ALL COME ON DIFFERENT SHIPS, BUT WE'RE IN THE SAME BOAT NOW."

– M. Luther King, Jr. (1929 – 1968)

After the discussion of major signalling pathways for development as well as the molecular mechanisms and network architectures keeping them on task, now is a good time to rewind the tape a bit. Let us go back to the "so simple a beginning" mentioned in the introduction. This is a single fertilized egg. How does the fertilization occur? How are the sexes and germ line cells specified? We begin with a disclaimer here that mechanisms of action and determinants are quite versatile across the animals. The steps of **germ cell** specification, the genetic **determinants of sex**, how organisms use **X-dosage compensation**, and even the very phenomenon of **fertilization** itself are often species dependent. Discussion will unroll first on that versatility, and we will wrap up with a focus on underlying conserved mechanisms such as initial **asymmetry** in the egg cell, **calcium waves** for **egg activation**, and the **maternal-to-zygotic transition**.

Versatility in germ cell specification

In sexual reproduction, the gametes combine to form the zygote which develops into the embryo. Gametes, **sperm** and **ovum**, differentiate from highly specialized cells called germ cells. Germ cells were kept aside at the very early stages of the parents' development when the embryo was being patterned into three embryonic germ layers, ectoderm, mesoderm, and endoderm. How does an embryo specify where its germ cells arise from? There are two ways which determine germ cell specification during early development: the germ plasm and fate induction.

The **germ plasm** is a pool of cytoplasmic determinants found in the egg. This is the most common way animals such as fish, birds, amphibians, and flies, specify where their germ cells arise from the very beginning. Most probably, this has convergently evolved many times across animal phyla due to its advantage keeping germ cells protected from dynamic signal changes happening in early development. In the pre-fertilized egg,

the egg cytoplasm is not evenly distributed, and its localization determines the region where germ cells, for either the ovum or sperm, will eventually come from. In mammals, however, germ cell specification arises from **fate induction**. Certain cells in early blastula are induced to specify as germ cells by external signal cues. We will elaborate more on this later.

Versatility in determinants of sex

Inheritance of **a sex-determining chromosome** is a good beginning to discuss how sex is determined. The XY system of mammals results in the Y-chromosome inheriting organism to develop into a male. That is to say, mammals with heterogeneity in their sex-chromosome inheritance develop into males (XY), whereas the embryo inheriting only X chromosomes (XX) become female. While sex-determining chromosome system is widespread among animals, it is not always the same as the XY system. For instance, we observe the ZW inheritance system in birds, where the heterogenous inheritance of the W chromosome results in the organisms being female (ZW), rather than male (ZZ). Although heterogeneity vs. homogeneity of sex-chromosomes can drive sex determination, this as well is not universal. Not all organisms rely only on the inheritance of a sex-determining chromosome and some instead rely on the copy number of the X chromosome. Although fruit flies encode a Y chromosome, the number of X chromosomes they possess ultimately decides whether the fruit fly will be male (X) or female (XX). Sex determination does not universally mean male or female in the animal kingdom either; *C. elegans*, a nematode worm, does not have males or females, but rather males (XO) or hermaphrodites (XX) being specified. Altogether, both copy number and heterogeneity of sex-related chromosomes can drive sex determination, but in ways that are unique to a given organism.

Further yet, some organisms do not rely on chromosomal inheritance at all for sex determination. **Environmental stimuli** can influence or drive sex determination as well. For many turtle and fish species, factors such as temperature and food availability during embryonic or juvenile development would result in male or female sex determination. A subset of fish species can “switch” their sex during adulthood with the appropriate environmental triggers.

Versatility in X-dosage compensation

Regardless of how sex determination takes place, it often results in a different copy number of X-chromosomes in male vs. female individuals. How do organisms properly regulate the expression of their genes given this disparity? Why doesn't a female mammal with twice the number of every single X-chromosome gene end up expressing those genes more? Different animals use different strategies to compensate for uneven distribution of shared chromosomes. Female mammals (XX) use **X-inactivation** to address the issue, where one of their X-chromosomes will be inactivated and maintained in the packed chromatin structure through epigenetic modifications. This is specifically done by coating the inactivated chromosome with RNAs, such that it is unavailable for transcription machinery. However, this as well is not a universal strategy; in male fruit flies (XO), we observe that they over-express the genes from their single X-chromosome to compensate for the **gene dosage** difference.

Types of fertilization

Versatile strategies are present in fertilization too. Many animals are **internally fertilized**, such as mammals, reptiles, birds, most amphibians, but some can also be **externally fertilized**, which is commonly seen in marine animals and some amphibians. However, some animals do not require fertilization at all for egg activation to take place. Through **parthenogenesis**, a subset of organisms can form an embryo in the absence of sperm (*e.g.*, aphids, whiptail lizards, sharks, Komodo dragons). Prevention of **polyspermy** (multiple entries of sperms into a single egg cell) is another important and versatile concept ensuring a proper beginning for the embryo, which will be discussed in greater detail later.

Sex determination and germ cell specification

What are the differences between somatic cells and germ cells? We know that when a fertilized egg, the zygote, starts developing many cells undergo differentiation and will form distinct tissue structures and exhibit specialized function. These cells are somatic cells, and they begin from totipotent embryonic cells – totipotent meaning they have the potential to be anything. Somatic cells undergo many rounds of mitotic division, generating distinct lineages during this differentiation and maturation process. However, germ cells

are maintained as a separate lineage from very early development, protected from committing into any other cell fate. This results in two process pipelines within the developing zygote. If the developing zygote is exposed to some mutagenic element (*e.g.,* X-ray irradiation, UV, or chemical exposure) that affects the somatic lineage, it will have no consequence on the germ line. That is to say, mutations accumulated in somatic cell lines cannot be inherited by the progeny of the adult organism. However, if a mutation were to arise in the germ cell line, it would then be passed to the organism's progeny.

Although germ cells do not result in the formation of complex tissue structures, they do still undergo specification. Rather than undergoing mitotic division, germ cells will carry out **meiosis**. In the first round of meiotic divisions, they will generate four, rather than two, copies of their chromosomal content and then undergo recombination. This event increases the genetic diversity of any subsequent progeny since the content is a mixture of the paternal and maternal lineages. However, depending on whether the germ cells develop into eggs or sperm, what happens following the first round of division differs.

The cells which eventually give rise to a differentiated ovum are referred to as **ooyctes**. Notably, in the prophase of meiotic division oocytes undergo meiotic arrest. So, the oocyte germ cells basically freeze all operations, until the organism becomes a sexually mature adult. A female organism would have a set count of oocytes prepared at the very early embryonic stages, which will be used up throughout their lifespan. On the other hand, the scenario for the germ cells giving rise to male gamete, the sperm, is a little different. There is a continuous stem cell differentiation for the sperm, which will undergo meiotic divisions to make the haploids which eventually, during fertilization, are going to combine with the haploid ovum of females.

Regardless of whether they develop into ovum or sperm, initial cells are referred to as **primordial germ cells (PGCs)**. We briefly went over how these PGCs are formed – either through the germ plasm inheritance or induction. We will go through these two scenarios in greater detail.

In some eggs, the germ plasm dictates where the PCGs will form. The localization of the germ plasm in the egg is especially important for this. For fruit flies, the germ plasm is localized to the pole of the oocytes in the follicle. Broadly, the pole germ cells in the oocyte will specify the whole germ cell lineage in these animals. Certain fish embryos, alternatively, have their germ plasm localized at one corner of the oocyte. Induction, alternatively, can also give rise to germ cell specification. Mammals have cells that are sensitive to the signals

that drive their differentiation into the primordial germ cell line. Rather than having a physically distinct compartment in the egg, the germ plasm, differential chemical signals play a role in mammalian PGC formation.

Despite these massive differences, there are conserved features between organisms that specify their germ cells through either induction or a germ plasm. The cells that specify into germ cells will, in a highly conserved manner, turn on specific genes (such as *Vasa*, *Nanos*, and *Piwi*) to drive the commitment of the germ cell lineage.

Oocyte to ovum transition

The egg is special hub for the development to take place at – for this we must first understand how an oocyte transitions into a mature ovum. Oocytes will always develop in the egg chamber of the follicle. Notably, follicles have specialized cells that support the growth of the oocyte; in flies, these cells are called **nurse cells** and in mammals they can be referred to as **corona radiata**. The specialized cells deposit components such as nutrients, proteins, and RNA signals. Given these come from the maternal lineage, we refer to them as maternally deposited components. As the oocyte is nurtured and supported, it begins to swell and becomes surrounded by a protective layer, *e.g.,* the vitelline coat or zona pellucida. Additionally, the mature ovum must also include the yolk, which plays a role post-fertilization by providing the nutrients and material resources that support embryonic growth.

Altogether, the structures necessary to support the growth of the oocyte to an ovum and allow for successful fertilization (such as the yolk) **requires the ovum to be large in size**. The size of the egg also allows for it to have polarity within its cytoplasmic contents. Within the egg, there are cytoplasmic factors that are segregated into distinct spatial domains. These factors can include proteins, chemical signals, or RNA, and are broadly referred to as either morphogenic or cytoplasmic determinants. The distribution of these determinants, and their polarity within the egg, enable asymmetric division in early fertilization. During asymmetric division, the cells dividing would at large be physically the same size but will inherit different concentrations of cytoplasmic or **morphogenic determinants**. This asymmetric division of material gives rise to two distinct daughter cells, not genetically, but in terms of their cytoplasmic content. To summarize, the

egg is significantly larger than the sperm because it is specially supported by the nurturing cells during the oocyte's development into the ovum. It needs a structure to accommodate the yolk which plays an integral role post-fertilization, and to maintain the distinct localized domains of different morphogenic determinants.

Despite the differences in size...

While both the egg and sperm are haploid cells that must fuse successfully to give rise to viable zygote, it is hard to immediately imagine how this can be done effectively given their size discrepancies. How is it that a sperm, which is a thousand times smaller than the egg, can fuse and combine its genetic content to its counterpart? First, we know that there are important protein complexes that mediate gamete interaction. These complexes facilitate both the recognition of the sperm and the egg by one another and ensure that the two stay bound during fertilization. Another logistical problem is deterring multiple sperms from binding to the same egg. This problem is called polyspermy and there are multiple mechanisms to avoid this. A third concept we will investigate is egg activation during fertilization by the sperm. These three concepts, the protein-protein complexes that mediate gamete recognition, the mechanisms that deter polyspermy, and egg activation, collectively overcome the logistical problems posed by the size differences between the two haploid cells.

The mechanisms that underly how the sperm and egg interact in the early stages of fertilization are just as variable as every other phenomenon we have highlighted, but nevertheless we see evolutionarily conserved genes implicated in this process. Of the few genes that have been elucidated (*Izumo*, *Spaca6*, *Tmem*), we will focus on the role of Izumo1. Murine studies identified that male infertility could arise from **gamete recognition** loss, and with *in vitro* and *in vivo* work found that Izumo1 on sperm mediated the recognition of the receptor Juno on the eggs of mice. The Izumo1-Juno interaction in the murine models drives gamete recognition, but what's notable is that Juno is absent from non-mammalian organisms, while Izumo is encoded in a variety of organisms, including fish. In fish, a protein known as Bouncer, with no evolutionary relatedness to Juno, binds to Izumo to mediate gamete recognition.

It should be surprising that two proteins, Juno and Bouncer, play the same biological role while lacking either sequence or structural similarities. Only recently has it come to light how these two proteins interact

with the same gamete recognition complex on sperm. Using artificial intelligence (AI)-guided protein complex predictions, Deneke *et al.* first aimed to characterize the gamete recognition complex of the fish sperm, which included Izumo1 and two other sperm proteins, Spaca6 and Tmem81. AI-model suggested that Izumo1, Tmem81, and Spaca6 form a heteromeric trimer. By tracking sperms with a fluorescent mitochondrial stain in live, they observed the sperms lacking Tmem81 failed to stably bind to the egg cell, confirming the AI-predicted interactions. What's notable is that Tmem81, Spaca6, and Izumo1 homologues are known to be encoded in fish, mice, and humans. This begets the question: how is the same protein complex recognize evolutionarily unrelated egg receptors? The machine-predicted model revealed the answer; while both Bouncer and Juno recognize the trimer, they bind to completely different interfaces of the complex. This result is unexpected. Intuitively, one would expect that the proteins playing a biologically conserved role would undergo comparable convergent selection to form ligand and receptor pairs which function similarly. However, Juno is an interesting mammalian evolutionary innovation. **It is important to reflect upon those findings to grasp how using both a molecular and developmental biology lens would prove beneficial to understand the innovation and conservation in evolution.**

Prevention of polyspermy and egg activation

Once the egg and sperm physically interact, what happens next? After a single sperm and the egg forms this recognition complex, fertilization begins. The sperm donates **centrioles** to the egg. These centrioles, made available to the egg cell from the sperm, lead to the mitotic cell divisions to begin following fertilization. So, the sperm brings not only its own haploid copy of the genes but also facilitate mitotic division by providing the necessary centrioles. This, importantly, is one of the reasons why polyspermy is so detrimental: too many centrioles from multiple sperm binding events leads to aberrant cell division. Rather than the zygote dividing first into two cells, it will divide into four and this abnormal division pattern eventually will cause the disintegration of cells and the death of the embryo. Thus, polyspermy compromises the viability of the embryo and must be prevented immediately following the fertilization.

Multiple ways are selected for diverse organisms to prevent **polyspermy**. In fish eggs, we see the use of mechanical restraint from polyspermy through the **micropyle**, which is a little hole in the fish egg where only

a single sperm cell can enter. In sea urchins, we observe that upon sperm binding, the egg changes the **electrical** potential of its membrane, and this switch in charge inhibits interaction with additional sperms. Finally, and most importantly, **calcium waves** in the fertilized egg protects against polyspermy and drives egg activation.

A Historic Overview: Discovery of Calcium Waves and Jaffe Family

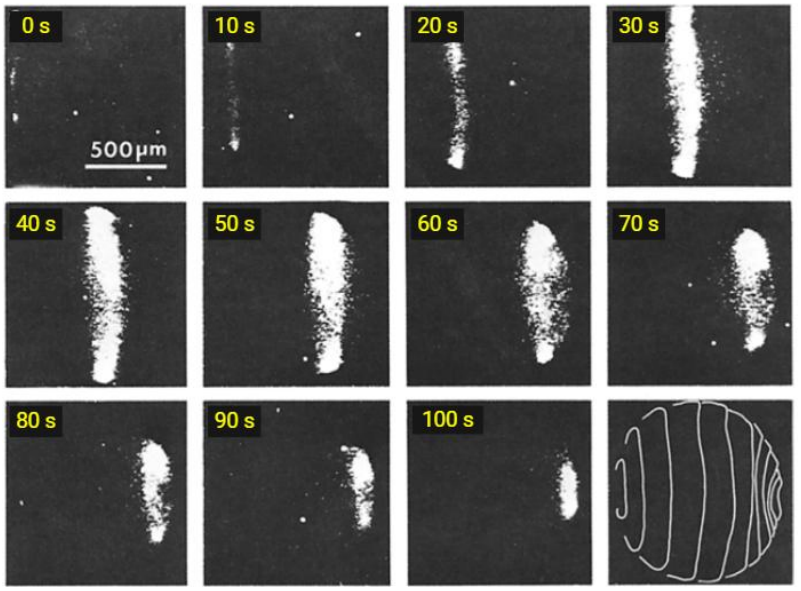

Figure 4.1 | (a) Lionel Jaffe (#23) together with scientists attending "Ionic currents in development" conference held at the Marine Biological Laboratory 3-6 September 1988.

(b) Time lapse of calcium wave traversing fertilized egg within a hundred seconds; from Egg #4 of experiments on 7 May 1976. Micropyle of medaka egg is aligned towards left centre.

The calcium waves following sperm recognition were first observed through time-lapse fluorescence microscopy in 1978 by Lionel Jaffe and colleagues. They used a calcium sensor

first discovered by the Nobel laurate Osamu Shimomura whose work on jellyfish fluorescence led not only to the discovery of the calcium-sensitive photoprotein, aequorin, but also of the renowned green fluorescent protein (GFP).

*Using aequorin as a calcium sensor, Jaffe et al. were able to directly visualize the calcium waves following fertilization of fish eggs. A technical hurdle in visualizing fertilization is understanding how to orient the egg and predicting where the sperm will bind to the egg. Lionel Jaffe was able to circumvent this by leveraging the previous knowledge that medaka fish eggs possess a micropyle, limiting the entry points of the sperm (**Figure 4.1**). This allowed the team to easily orient and position the egg to observe the calcium gradients resulting from fertilization.*

Around the same time Lionel Jaffe was providing explicit evidence for the calcium waves that follow fertilization of the egg, his daughter Laurinda Jaffe, a graduate student across the continent at UCLA with Susumu Hagiwara, identified the mechanism of how sea urchins block polyspermy. Laurinda Jaffe found that, in addition to the slow block by activating calcium waves, the membrane potential of the sea urchin eggs switch on electrically. This membrane potential switch acts as a fast block preventing egg entry of additional sperms by electrostatic repulsion.

Calcium waves are observed typically following sperm entry. We see that a wave of calcium begins at the site of sperm entry then dissipates throughout the cell; how many wave events take place depends on the

organism studied. For sea urchins, we observe a single wave event, while in humans there are at least four. Although we know that successive waves go through the embryo, we know little about why this is so in mammals. In diverse species (frogs, chicks, humans, mice, *etc.*) calcium waves only take place following fertilization. A notable exception is the fruit fly embryo, where the egg experiences calcium waves and activation prior to leaving the follicle and fertilization. Again, biology always has exceptions, but the importance of calcium waves during fertilization is a key conserved feature.

As the calcium spreads through the cell, it reaches the **cortical granules**, which are vesicle structures produced by the Golgi apparatus that surround the egg. Once the calcium comes in, the cortical granules explode and release their contents, which include glucose, aminoglycans, proteoglycans, and specialized enzymes. This release has a two-fold effect: blocking of polyspermy and the activation of the egg. The contents from the cortical granules are released in the spacing between the vitelline membrane and the egg surface, which enlarges the **fertilization membrane**. Enzymes released following granule lysis crosslink the vitelline coat such that it becomes tight-knit and impenetrable. Furthermore, the lysis of the granules also instigates cleavage of the egg-receptors that mediate gamete recognition. Altogether, further sperms can no longer enter due to the crosslinking of the space within the vitelline coat and the absence of their cognate receptor.

In summary, the sperm enters the egg through specialized gamete recognition receptor complexes (such as the Izumo1, Spaca6, Tmem81 trimer) which activates the phospholipase-C activated IP_3 production. This pathway stimulates the cumulative release of calcium through the fertilized egg, which then generates waves of calcium spanning the cell. This gradient reaches the cortical granules on the surface of the egg, which then rupture and release their contents. The release of cortical granule contents stimulates the cleavage of sperm-recognition complexes, the crosslinking of the vitelline coat, and the expansion of the fertilization membrane. Those events serve as a slow block to polyspermy and this whole process is also known as **egg activation**. Once the egg is activated, the initial entry of the sperm restores mitotic division by the sperm-mediated delivery of the centrioles. So, the fertilized egg leaves its years-long period of mitotic arrest and begins dividing again.

One last mechanism that limits polyspermy is the practical issue of sperm **migration**. Sperms have a long way to travel to reach the egg, which limits their ability to reach the intended destination in multitudes. This

holds true both for internal and external fertilization, the latter imposing this issue due to the three-dimensional nature of the travel in a freely diffusive environment. This practical problem for the sperm is also a solution to prevent the polyspermy by limiting how many sperms can reach the egg.

FURTHER READING

The mammalian germline as a pluripotency cycle. H. G. Leitch, A. Smith, **Development**, 2013. DOI: 10.1242/dev.091603.

A conserved fertilization complex bridges sperm and egg in vertebrates. V. E. Deneke *et al.*, **Cell**, 2024. DOI: 10.1016/j.cell.2024.09.035.

A free calcium wave traverses the activating egg of the medaka, Oryzias latipes. J. C. Gilkey, L. F. Jaffe *et al.*, **Journal of Cell Biology**, 1978. DOI: 10.1083/jcb.76.2.448.

Chapter 5: FROM FERTILIZATION TO GASTRULATION

"TO OUR LIMITED INTELLIGENCE, IT WOULD SEEM A SIMPLE TASK TO DIVIDE A NUCLEUS INTO EQUAL PARTS. THE CELL, MANIFESTLY, ENTERTAINS A VERY DIFFERENT OPINION."

– E. B. Wilson, 1923.

Once the egg is fertilized, the sperm donates its centrioles and allows for the zygote to undergo mitotic divisions. Early series of mitotic divisions, the cleavage events, result in formation of many smaller cells. An important facet of cleavage is **asymmetric** division, such that the cells formed vary in content, *e.g.,* proteins, chemical gradients, and RNA transcripts, even if they are genetically identical. The **polarization** of the oocyte and activated egg drives this eventual asymmetry of maternally deposited morphogenetic factors. This polarity leads to the formation of the first axis we define for the fertilized egg: **the animal-vegetal axis**. The animal pole of this axis is the site for the formation of the embryo proper, the true embryonic mass, while the vegetal pole is the site of the yolk, which provides important nutrients during development but does not contribute to the embryo directly, hence 'vegetal'. Not all animals have the same size or proportion of **yolk**. Animals with externally developing embryos, such as fish, frogs, and birds, typically have larger yolks. A famous example is chicken. However, for mammals, such as humans, the yolk is almost non-existent. The embryo implants into the uterine wall and directly gets nutrients and developmental support from the mother. Sea urchins, although externally developing, also lack much yolk content. They instead quickly develop into their larval stage (pluteus) and suck nutrients directly out of the sea water. So, while in both cases, the organisms exhibit a reduced yolk size, they obtain their required nutrients during development through quite different strategies.

Case Study: From Maternal Factors to Gastrulating Germ Layers

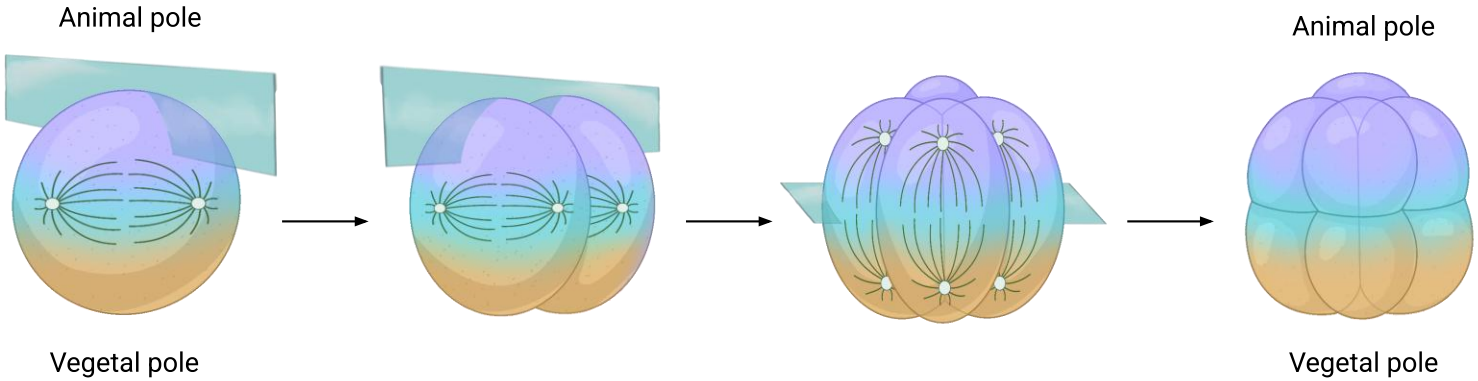

Figure 5.1 | Asymmetric cell division in sea urchin embryos.

Let us elaborate on the sea urchin case to understand how the early cleavage results in an asymmetric multicellular mass, the blastula, which will give rise to distinct germ layers. As they give rise to blastula, the earliest cells during cleavage are referred to as blastomeres.

*How does the mitotic divisions of the blastomeres drive structural asymmetry? The animal vegetal axis arises from the polarity of the maternally deposited determinants, already maintained by the oocyte (**Figure 5.1**). However, the first and second rounds of mitotic division in the sea urchin occur perpendicular to the animal-vegetal axis, such that the four resulting daughter cells exhibit same polarity with the starting single-cell. It is only in the third round of division, which takes place parallel to the animal-vegetal axis, that we observe **asymmetric division**. As the split occurs across the animal-vegetal axis, two distinct sets of*

*blastomeres arise at the eight-cell stage: four which have inherited the morphogenetic determinants from the animal pole, and the other four with the determinants concentrated to the vegetal pole (**Figure 5.1**). This asymmetry has important consequences for subsequent division events. The most pertinent example is seen through assessing the orientation of the mitotic spindles between the "animal" and the "vegetal" sets of cells. In the animal cells, the mitotic spindles are situated perpendicular to the original animal-vegetal axis. The vegetal cells, however, have mitotic spindles not only running parallel to the animal-vegetal axis but are also not evenly situated within each cell, which ultimately results in the formation of differentially sized daughter cells after the fourth division event, at the 16-cell stage. The larger cells, macromeres, are situated closer to the animal pole and are "on-top" of the smaller cells, the micromeres, which rest at the vegetal pole.*

The Wnt signal gradient

Interesting enough, following the 32-cell stage, the micromeres completely halt their mitotic divisions, resulting in the formation of a 60-cell mass, rather than 64 cells. Additionally, micromeres seem to be the first site of Wnt-signalling activation during embryogenesis. This is because the dishevelled protein, the activator of the Wnt signalling cascade, is a morphogenetic determinant concentrated to the vegetal pole. As such, it is one of the factors inherited during the early cleavage events by the micromeres. Dishevelled proteins prevent formation of destruction complex in cytoplasm which phosphorylates β-Catenin and targets them for degradation. Undegraded β-Catenin can then localize to the nucleus to activate gene expression. During early blastula, β-Catenin enters the nucleus of micromeres and neighbouring veg2 cells to stimulate the Wnt signalling cascade. The resulting cascade

establishes a gradient along the embryo from the vegetal pole to the animal pole: this ultimately drives the specialization of cells to give rise to different germ layers.

Correlation vs. causation in development

What does it mean for the Wnt pathway to be important in early patterning of sea urchin embryos? How do we know that the resulting signalling cascade is an instructive signal during development? Scientists perturbed the ability of β-Catenin to enter the nucleus to either halt or overstimulate Wnt-activation. The Wnt signalling cascade can be overactivated by lithium chloride treatment. Lithium chloride acts to block GSK3-β, an enzyme inhibiting β-Catenin's translocation into the nucleus. This results in β-Catenin entry into the nucleus at every cell in the embryo, activating and over-saturating the Wnt cascade. Phenotypic outcome of LiCl treatment is abnormal germ layer specification – the resulting embryos have no ectoderm, and the larvae only develops from the mesoderm and a bit of an endoderm. To deplete β-Catenin from the nucleus on the other hand, scientists overexpressed a cytoplasmic Cadherin. Binding the β-Catenin, this sequestered the β-Catenin away from nucleus. In the absence of Wnt-activation, the embryo forms no internal layers, and the larva is only comprised of the ectoderm. The perturbations to the Wnt-signalling cascade provide compelling evidence that its graded signal from vegetal to animal pole plays an important role in the formation of the germ layers during embryogenesis.

Cleavage

Cells are getting smaller and smaller as the mitotic divisions proceed in an early embryo. This is cleavage. Even though "cleavage" results in a similar outcome, how it is carried out varies wildly between organisms. As E.B. Wilson noted in 1923, *"to our limited intelligence, it would seem a simple task to divide a nucleus into equal parts. The cell, manifestly, entertains a very different opinion".*

How is it that, in all these diverse ways, cells are dividing without changing their total mass? To address this, a refresher for **the cell cycle** concept is beneficial. First discovered by frog embryologists in the 1800s, "the cell cycle" is a familiar face for the biology student. During the cell cycle, ordinary cells in an adult stay in inter-, the non-mitotic, phase, for a long time. When they divide, the mitotic phase happens quite quickly. However, beginning with the zygote, the cells skip interphase and are not making more material before entering mitosis. These cellular division events result in more cells, but not more mass.

While the observation that blastomeres skip interphase explains how we see no increase in total mass, it does not answer why this is happening. Cyclin B1 is central for the mitotic phase and its level decreases after mitosis. What does Cyclin B1 do during mitosis? Cyclin acts to activate the Cyclin dependent kinase (CDK); CDK and Cyclin B1 are referred to as the mitotic or maturation promoting factors. Together, they act to promote division. More specifically, CDK prepares the cell for mitosis by phosphorylating protein targets and is estimated to have between 50 to 700 unique protein substrates. An important one is the anaphase promoting complex (APC); APC acts as a **negative feedback** stimulus, which allows for the oscillations to take place. Negative feedback oscillators were previously discussed, for example with the Notch feedback loop.

The oscillations mediated by APC, however, are not robust and eventually observe a decay. However, we see that coupling this APC-mediated negative feedback with **positive feedback systems** increases the **robustness** of the oscillations. Here in the fast cycles of cleavage events, two distinct positive feedback are at play: the direct positive feedback of Cdc25 and the double negative feedback (which becomes positive) of Cyclin B1. What is the effect of changing the strength of those positive feedback systems? Well, as one changes the strength of positive feedback, they can tune the **frequency** of the cell cycle. Besides minimizing decay in oscillations, "positive feedback loops" additionally allows to control the speed of the cell cycle. The oscillations

initiated by the APC negative feedback exhibit less decay when coupled with positive feedback from Cdc25 and Cyclin B1. Notably, the zygote is uniquely primed with these positive feedback factors Cdc25 and Cyclin B1, which are maternally deposited cytoplasmic determinants. The materials to divide are provided by the yolk or the environment (ex. uterine lining). The collective events of the zygote being primed with the positive feedback factors and having external nutrient resources allow it to bypass interphase and undergo a series of rapid mitotic divisions without increasing its overall volume.

Why is this positive feedback so uniquely important? CDK activation in the centrosome triggers the cell cycle. When CDK is active, a cascade resulting in the activation of other CDKs that phosphorylate necessary targets throughout the cell is initiated. This signal relay relies heavily on the diffusion of CDK in a regular adult cell; however, an egg is significantly larger, such that diffusion alone would fail to deliver the task. Just as we saw for the calcium waves of egg activation, the egg depends on the positive feedback to generate a chemical wave that is carried throughout the cell volume. Positive feedback mediated by factors such as Cdc25, drives the chemical wave of activation along the embryo that serves to **synchronized** cell divisions.

Cell division axes of early cleavage events are determined according to egg/yolk geometry. In the case of fish, there is a big yolk, and the early divisions happen in a single plane, which we call discoidal cleavage. This is similarly observed in chick embryos. However, in sea urchins and mammals, we do not have a separate animal and yolk set dividing and instead see division throughout the egg. This form of division is called holoblastic cleavage. In fruit flies, the yolk is in the middle, and the cells divide around its edges. The cleavage axis forms different in different eggs, and we see varying degrees of symmetry/asymmetry. Our conclusion should be that **the plane of division**, which ultimately affects the geometry of the blastula and the embryo, is greatly influenced by the size and relative position of the yolk.

Maternally deposited factors during gastrulation

The maternally deposited factors play important roles not only in early division events. We see that in fruit flies maternally deposited RNA transcripts play a key role in the formation of the larval body during gastrulation. Beginning with early fertilization, in the first thirty minutes following fusion of the egg and sperm, there is a single nucleus in the cytoplasm of the zygote. A cytoplasmic medium that surrounds the yolk, which

is referred to as **the syncytial layer**, allows for the nuclei alone to continue dividing the cytoplasm in the fruit fly zygote. It is only after two to three hours of development that we observe the cellularization of the syncytial blastoderm, and these many nuclei become encased in their own independent cell membrane; a process quite specific to certain insect species.

What is the result of these many nuclei sharing a cytoplasm during early development? It all comes back to cytoplasmic factors. Alongside protein determinants, like Cdk1 and Cyclin B1, which we previously discussed, there are numerous other RNA transcripts and proteins which are deposited in the egg in a localized fashion. These determinants can more immediately diffuse to the dividing nuclei when they share a cytoplasm. Inheritance of these determinants plays important roles in gastrulation and shapes the pattern of the anterior-posterior axis of the fertilized embryo, as we will see in detail in upcoming chapters. Because these determinants affect the morphogenesis of the organism, they are also referred to as **morphogenetic determinants**. We will briefly introduce three maternally deposited RNA transcripts (mRNA) that act as morphogenetic determinants in fruit fly embryos:

1. **Anterior patterning.** The mRNA encoded from the *bicoid* maternal allele is deposited on the anterior side of the egg. Absence or mutation of *bicoid* results in the loss of head specification.

2. **Posterior patterning.** We see that mRNA encoded from the *nanos* maternal allele is deposited on the posterior side of the egg. Absence or mutation of *nanos* results in the loss of the tail/posterior specifications, ultimately manifesting in an underdeveloped and truncated abdomen.

3. **Terminal patterning.** The transcripts encoded from the *torso* maternal allele is deposited to each of the terminal ends of the *Drosophila* egg. Maternal zygotic mutations cause aberrant development of both the posterior and anterior ends.

The maternally deposited morphogenetic factors define the early polarization of the cell, the formation of the animal-vegetal axis, influence cleavage, and mediate specifications during gastrulation. Different organisms mediate these factors as needed to furnish the early embryo in different ways. We observed how positive feedback systems amplify chemical waves over long distances and how fruit flies leverage a shared

cytoplasm to effectively diffuse even RNA transcripts. However, the zygote will eventually come to no longer need the maternal factors during its development. We call this stage the maternal-zygote transition.

Maternal-zygotic transition (MZT)

The oocyte sets the scene for early protein activity and expression, which drives the patterns we initially see. However, in all organisms, they eventually transition away from relying on the maternally deposited factors and begin instead making use of their own genetic content. In yolk-rich embryo animals, the MZT is called the mid-blastula transition since it takes place in the middle of the blastula stage. For those species, we observe during mid-blastula stage that the cell cycle slows down and that division becomes less synchronous. In fruit flies, MZT occurs in the late blastula (post-cellularization) stage. In *C. elegans*, it is much earlier – at the mere four-cell stage, and even sooner for sea urchin and mice embryos. We do not currently have enough data to suggest how similar human MZT is to other animals, such as mice. The important takeaway is that as the cleavage events give rise to the blastula, the zygote will need to activate its genome. MZT is a tiered process. The ordering broadly follows as such: the maternal materials are degraded, the zygotic materials are activated, and then the full dependence on the zygotic genome follows. We will emphasize four main takeaways from this extremely broad topic of maternal-zygotic transition:

1. **The destabilization of maternal transcripts.** The maternal RNAs are deposited at specific positions in the oocyte and stably kept there for a long time. But, following activation, they will need to be degraded within the measure of hours. There are molecular machineries dedicated to the destruction of the maternal mRNAs, but there are different biological models as to why this needs to happen. One idea is that of RNA level compensation. It is thought that when you have activation of the zygotic genome, your RNA levels might get too high and resemble an overexpression phenotype. Consequently, one can think degradation of maternal transcripts is a precaution step before zygotic alleles being transcribed. Another line of thought proposes that the degradation of the maternal RNAs might be instructive for the activation of the zygotic genome.

2. **The nuclear-cytoplasmic ratio.** During cleavage, we know that the cells are getting smaller and smaller, but that their nuclei do not change much in size. Extensive literature shows that ratio of cytoplasm to nuclei may be one of the most important factors triggering the zygotic genome activation.

3. **Epigenetic modifications.** We know that the chromatins of the zygote are in a closed (heterochromatin) state with a tight control of epigenetic modifications. During MZT, the chromatins enter an open (euchromatin) state and get ready to be read for expression.

4. **Totipotency to pluripotency.** The MZT funnels the possibility of differentiation for cells once the zygotic genome is activated, such that they can no longer develop into any other cell type.

FURTHER READING

Nuclear beta-catenin is required to specify vegetal cell fates in the sea urchin embryo. C. Y. Logan *et al.*, **Development**, 1999. DOI: 10.1242/dev.126.2.345.

Mitotic trigger waves and the spatial coordination of the Xenopus cell cycle. J. B. Chang, J. E. Ferrell Jr., **Nature**, 2013. DOI: 10.1038/nature12321.

The maternal-to-zygotic transition: a play in two acts. W. Tadros, H. D. Lipshitz, **Development**, 2009. DOI: 10.1242/dev.033183.

Chapter 6: PRINCIPLES OF MORPHOGENESIS

"CELL AND TISSUE, SHELL AND BONE, LEAF AND FLOWER, ARE SO MANY PORTIONS OF MATTER, AND IT IS IN OBEDIENCE TO THE LAWS OF PHYSICS THAT THEIR PARTICLES HAVE BEEN MOVED, MOULDED AND CONFORMED. THEY ARE NO EXCEPTIONS TO THE RULE THAT "ἀεὶ ὁ θεὸς γεωμετρεῖ". THEIR PROBLEMS OF FORM ARE IN THE FIRST INSTANCE MATHEMATICAL PROBLEMS, THEIR PROBLEMS OF GROWTH ARE ESSENTIALLY PHYSICAL PROBLEMS, AND THE MORPHOLOGIST IS, IPSO FACTO, A STUDENT OF PHYSICAL SCIENCE."

– D'Arcy W. Thompson, 1917.

We have previously noted that cells are the builders of the embryo and emphasized the four cellular activities that make this possible: migration, multiplication, signalling, and change in character. In the roller-coaster of embryonic development, our next track is 'morphogenesis'. Morpho-genesis is a combinatorial word coined in 19th century by combining "μορφή" for "shape, beauty, form", and "γένεσις" for "birth, creation, formation" from Ancient Greek. Two main characters of cellular action relevant for 'morphogenesis' are 'migration' and 'multiplication'. Through a choreography of moving into places and multiplying in certain times, many cells from a single zygote mold into the embryonic form. Let us aim to understand the constructive repertoire of a singular cell first to comprehend how those building blocks of the embryo work as its builders.

Morphogenesis for a single cell

In the most simplistic form, two structures give rise to the shape of a cell: microtubules and actin filaments (**Figure 6.1**). A **microtubule** is a spiraling cylindrical molecule formed by two protein subunits, α-tubulin, and β-tubulin. The protein complex has very specified diameter based on the sizes and coordination of tubulins that interact with each other to make the cylindrical structure. These structures have two ends, a **negative** slow growing end mostly anchored to microtubule organising centre (**MTOC**) or **centrosome** of the cell, and a **positive end**, where the polymerization and growth are observed. The positive end usually grows towards the membrane of the cell and nucleotide exchange (GDP to GTP) enables polymerization from this tip. Once bound

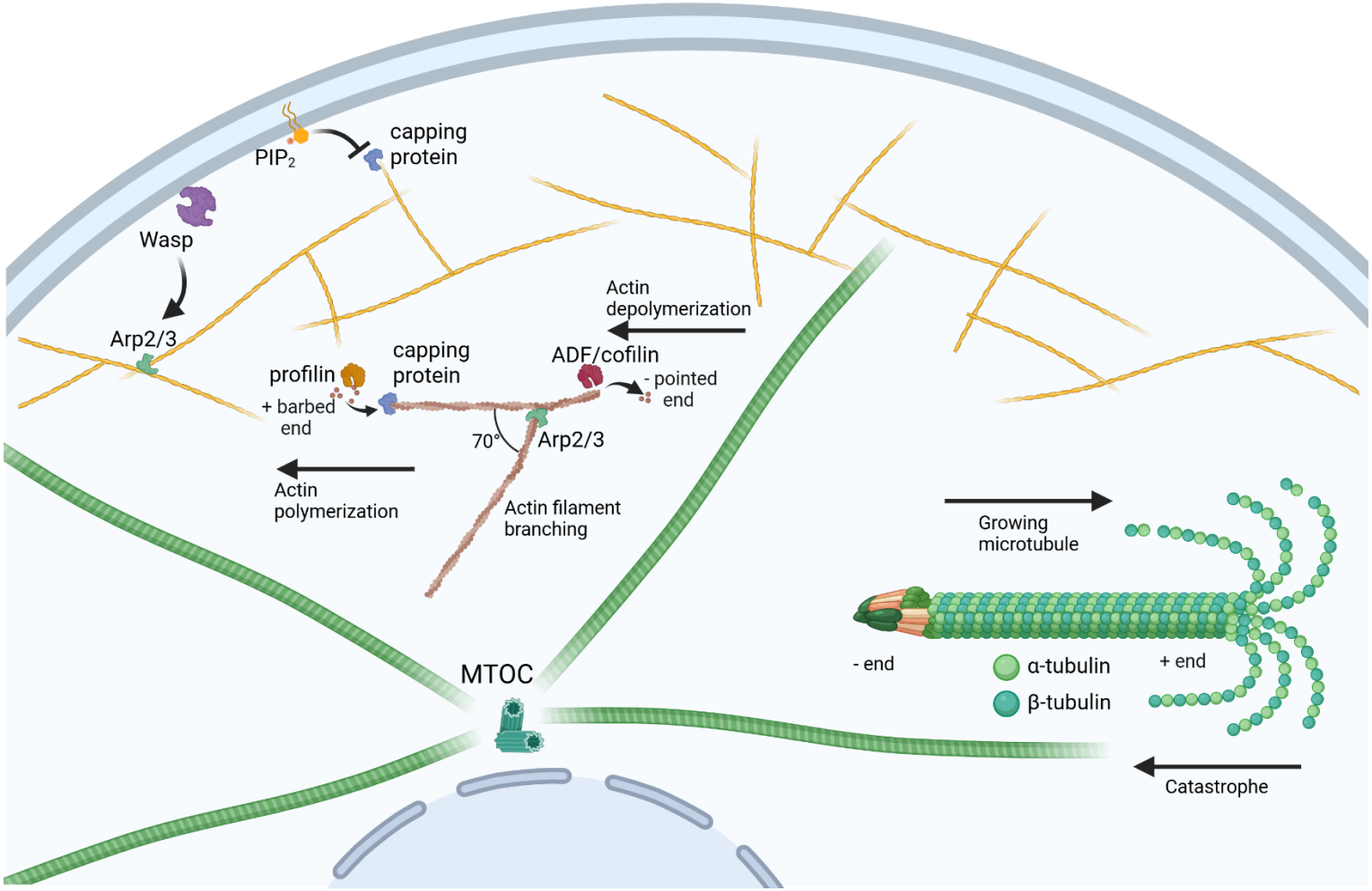

Figure 6.1 | Organisation of cytoskeleton and molecular basis for its assembly.

GTP of tubulin dimer gets hydrolyzed to GDP deeming the microtubule structurally very unstable. If left alone, microtubule rips apart in an event known as **catastrophe**.

Microtubules are very rigid structures and are 300× stiffer than the actin filaments. Their rigidity is important in many developmental processes. Broadly, microtubules support the cell shape, drive mitotic spindle formation, and help define the cell division axis.

Actin filaments are composed of polymerizing actin monomers. Actin molecules that are bound to ATP can treadmill towards the growing, or **barbed**, end of the filament. Treadmilling helps maintain the actin pool within the cytoplasm, as cells require actin filaments to impart flexibility to their cellular structure.

Actin filaments tend to grow towards the direction of the membrane. What determines the length of an actin filament? A filament is polymerized with the assistance of profilin proteins, which associate with ATP-bound actin subunits and facilitate their incorporation into the barbed end of the polymer. Two additional factors can alter the rate of synthesis: capping and depolymerization. When a **capping protein** binds to the barbed end of the filament its stops actin filament growth. Additionally, factors like actin depolarization filament (ADF) and Cofilin proteins mediate the disassembly or depolymerization of actin. Rates of depolymerization, based on the amount of **ADF/Cofilin** bound, help modulate the length of the actin filament. Put simply, if the filament is depolymerizing quickly (through ADF/Cofilin activity) one ends up with a short fragment but if the filament depolymerizes slower, it will get longer. By controlling ADF/Cofilin levels, a cell can locally control the length of the actin filaments.

Actin polymerization happens towards the membrane, but how is it possible that filament growth is maintained once the membrane is reached? Here, thermodynamics comes into play. Because of the ambient temperature reactions taking place, the cell membrane and actin filaments will be stochastically "flexible" and be more jiggly-wiggly in terms of stability. This temperature-dependent random motion of molecules is referred to as **Brownian motion**, and we know that with certain probability, changes will take place in accordance with the energy barrier associated with that change. Occasionally, the membrane will fluctuate in space, and the actin filaments will swing and stretch; collectively referred to as the **Brownian ratchet**

mechanism. This stochastic movement causes a new filament to be available for polymerization and enables the filament to keep growing as it pushes against the membrane.

Actin does not grow unidirectionally, as we observe for microtubules. There are specialized protein complexes that enable actin filament bifurcation, where a new actin filament can grow from an older one, much like a branching tree. The **Arp2/3** protein plays a significant role in complex formation. In simplistic terms, Arp2/3 will bind to an "old" actin filament and form a larger protein complex that allows the filament to grow in a different direction. Fascinatingly, this protein complex will dictate the new filaments always branch 70 degrees away from the starting structure. As an aside, let us rewind back to the Brownian rachet mechanism. If the filaments were to push against the membrane at 90 degrees, it would be harder to add new actin to that filament. If you were going to add a branch at zero degrees it would be much easier to polymerize, but you would not get any amount of membrane pushing. The angle (70 degrees) set by the molecular nature of actin filament branching seems to be a "sweet spot" that allows for membrane pushing while still having a decent probability of adding actin to growing filaments.

How does a cell control when the filament branching happens? Arp2/3 mediated actin branching is activated by the **WASP** membrane-bound protein. The localization of WASP to the membrane ensures that, an actin filament closer to the cell's membrane experience more Arp2/3-mediated bifurcation. Likewise, the further a filament is from the membrane the less likely it will be that bifurcation takes place. The localization of WASP ensures that filament branching is restricted towards the membrane.

An additional layer of cellular control to ensure filament polymerization is enriched near the membrane is PIP2-mediated inactivation of the capping proteins. **PIP2**, as a signalling lipid molecule, is localized to the membrane and inhibits the ability of capping proteins to bind to the barbed end of actin filaments (**Figure 6.1**). This means that the filaments closer to the membrane will not be capped as easily and will continue to grow near the membrane. Conversely, if you have an actin filament growing in the "wrong" direction, away from the membrane, it is more likely to interact with an active capping protein. This ensures that the cell does not waste its actin filaments, and all growth drives pushing of the membrane.

These collective strategies deter backwards filament polymerization, and the nanoscale chemistry helps leading edge directed cell migration. During movement, the leading edge of the cell is filled with growing

filaments and forms a cellular structure known as lamellipodium. This structure allows cells to crawl and **migrate**.

How is the cell shape maintained?

The actin filaments and microtubules collectively give a cell its shape. Their collaborative effects can be summed into the **tensegrity model**. This model, which is derived from principles of architecture, can be described as tension-dependent integrity. Consider a simpler example: you have two pencils and an elastic band. If you cross the pencils such that they make an "X", with the midpoints in contact, and wrap the elastic band around the four edges, you can maintain this structure only through a) the integrity provided by the pencils rigidity and b) the tension imparted by the elastic band's flexibility. Similarly, a cell's shape is defined by the rigid microtubules while the tension is maintained by the cell's flexible actin filaments.

Epithelial cell polarity

In the earliest division events, we have described polarity in a few different contexts. We know polarity can refer to the different contents inherited by daughter cells, such that between different cells there are differences in cytoplasmic content, or that it can refer to the polar intracellular distribution of cytoplasmic content within a single cell. We will be solely using the latter definition in the following section.

Knocking on the doors of embryo returning a quick cell biology recap, let us consider again the early cell divisions. These rapidly dividing cells, which are in close contact with one another, are categorized as epithelial cells, and they form epithelial sheets important for gastrulation. Epithelia themselves are individually polarized, possessing a **basal** membrane ("bottom") and an **apical** membrane ("top"). At large, this polarity is maintained by the microtubules; the negative (not growing side) is fixed to microtubule organizational centres (MTOC) that are on the apical side of the membrane. The positive end of the microtubule grows towards the basal membrane. Localized to the apical membrane is also a mesh network of actin filaments which form a sort of **contractile belt** that can squeeze and expand, which is important for changing the shape of the epithelial cell. Epithelia are characterized by their cell-to-cell contact with neighbouring cells. There are four ways in which

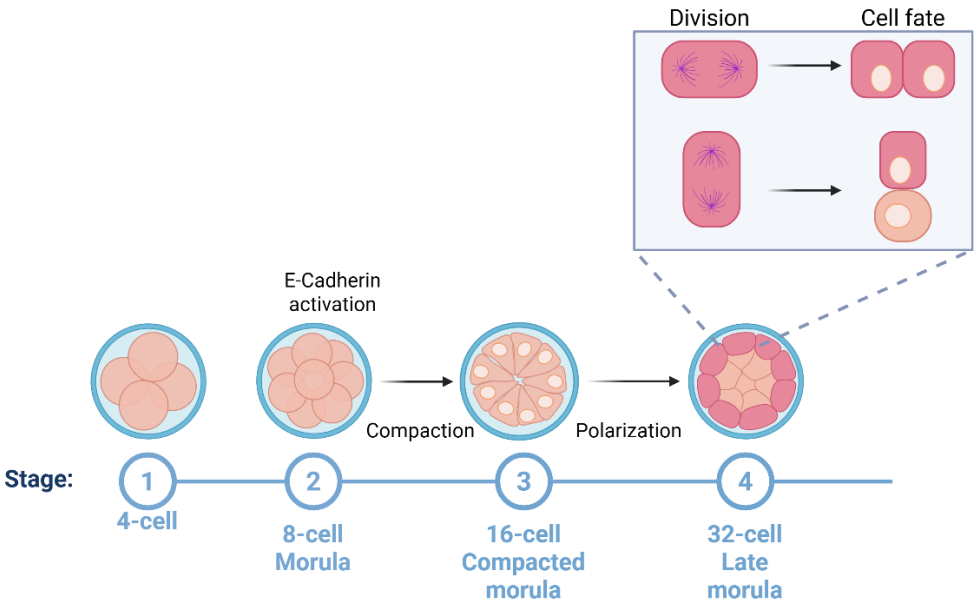

Figure 6.2 | Early cell divisions and patterning of mammalian morula.

cell-to-cell connections are maintained, but for the purposes of this course we will focus on the **adherens junctions** established by Cadherin proteins.

In a mammalian embryo, at the third division event, E-Cadherin is activated. In turn, 16-cells undergo compaction and become polarized, so that outer side of the membranes are apical while the inner sides are basal (**Figure 6.2**). Why is this important? Following this cellular polarization, at the fifth division event we end up with two distinct types of cells: the **trophoblasts**, which is the surrounding epithelial layer, and the **inner cell mass**. The inner cell mass eventually gives rise to the embryo, while the trophoblasts support the inner cell mass itself. As every single epithelial cell is polarized with a basal and apical side, how does a cell know whether it will contribute to the inner cell mass or the trophoblast? This cell fate is determined by the axis of cell division.

If the cell divides along the plane, then both daughter cells will stay in the trophoblast. However, if their division is perpendicular to the axis, then they will get kicked out of the epithelial layer and join the inner cell mass (**Figure 6.2**). It may make one awe-struck to think that, whether a cell became part of you and me, us humans learning, discovering, and thinking about those concepts, were determined by the **stochastic** alignment of division axes during those early stages of development.

Another example of how cell morphology can give rise to tissue morphologies is something we have previously examined while discussing 'induction' concept for the first time: formation of the lens cup of the eye from **lens placode**. Following induction from the neural ectoderm, the epithelial ectoderm tissue will form the lens cup. We will now dive into how cell morphology drives the specification we observe under the influence of this induction. When the cells of the epithelial ectoderm receive the signal, they change their cell shape and expand. This **expansion** is driven by the microtubules, as their rigidity provides the structural basis for the cell's shape. The cells elongate through microtubule expansion and then actin filaments begin to form stretchy contractile rings that convert the cell from a columnar to a wedge shape. This collective shape change in the epithelial cells results in the **invagination** of the epithelial sheet.

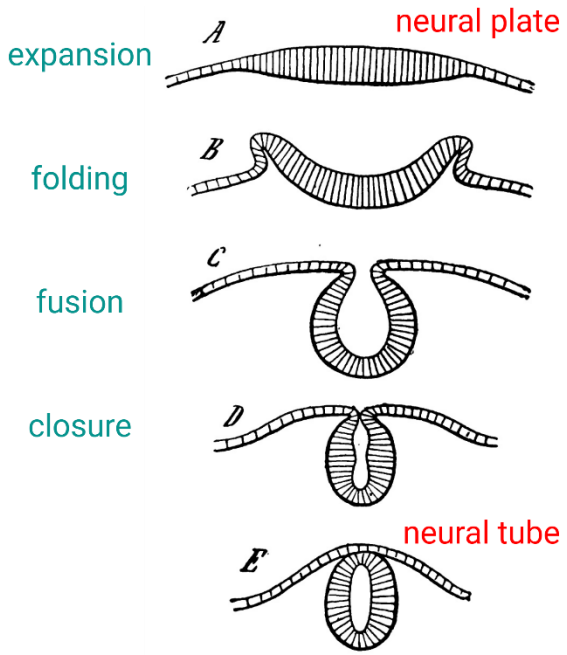

Figure 6.3 | Morphogenetic description of neural tube formation. Adapted from J.W. Jenkinson, Experimental Embryology, 1905.

A third example is morphogenesis of the **neural tube** (**Figure 6.3**). This tissue originates from a single layer of ectodermal epithelial cells called the neural plate. The neural plate begins the formation of the neural tube structure following the recognition of induction signals released by the notochord underneath. Again, following signal recognition, we observe **expansion** mediated by microtubule elongation, followed with actin filament-driven **folding**. Eventually, actin filament contraction and stretching result in **fusion** of the tube. During the folding process, the cells are pushed together, and this **closure** leads to the closing of the sheet structure and the formation of the neural tube. Similar to the formation of the inner cell mass earlier in development, the cell division axis

drives the correct positioning to enable the development of the tissue structure. The cells of the neural epithelium, at this stage, are dividing along a single perpendicular axis while folding is taking place. This causes a directional push that shifts the epithelial sheet towards the middle. If the divisions were randomly oriented, we would observe expansion, but no specific directional movement. Controlling the directionality of cell division drives the push, and it is only because the cells are dividing along a shared axis that we see fusion and closure to complete the process.

All those three examples teach us that individual cell morphology gives rise to tissue morphogenesis during development. Cell division axes control is a strategy to alter resulting tissue shapes, and microtubule-actin coordination can make tubes out of sheets, like pieces of origami!

FURTHER READING

Roles of the actin cytoskeleton and cell adhesion in tissue morphogenesis. D. N. Clarke, A. C. Martin, **Current Biology**, 2021. DOI: 10.1016/j.cub.2021.03.031.

Biochemical and mechanical regulation of actin dynamics. P. Lappalainen *et al.*, **Nature Reviews Molecular Cell Biology**, 2022. DOI: 10.1038/s41580-022-00508-4.

Molecular mechanisms of cell shape changes that contribute to vertebrate neural tube closure. M. Suzuki *et al.*, **Development, Growth & Differentiation**, 2012. DOI: 10.1111/j.1440-169X.2012.01346.x.

Chapter 7: GASTRULATION

"IT IS NOT BIRTH, MARRIAGE OR DEATH, BUT GASTRULATION WHICH IS TRULY THE MOST IMPORTANT TIME IN YOUR LIFE."

– L. Wolpert, 1979.

As discussed in the previous chapter, an epithelial cell is polarized between their basal and apical membranes, with the apical side having a thick actin meshwork. In this polarization, we usually see the nucleus of the cell closer to the basal side. Although the nucleus is not fixed to the basal side, that seems to be its most stable position. In future lectures, we will observe that the nuclei within a developing epithelial sheet will actually travel between the basal and apical ends; coming to the apical side of the cell, they will become more prone to division. A second polarity to note is the molecular one, for the microtubules and the actin filaments. Those structures are polar in the sense that for instance actin monomers polymerize at the barbed (positive) end of the filament and depolymerize from the pointed (negative) end. Likewise, negative end of microtubules merge at the microtubule organizing centre (MTOC) which is typically localized pretty close to the nucleus for a regular cell, however, near the apical surface for an epithelium. One last note for epithelial architecture is the side membranes which form cell-to-cell junctions. For the sake of simplicity, our focus in the upcoming discussions will mostly be on adherens junctions and homophilic adhesion of transmembrane Cadherin molecules forming those connections.

Let us dive deeper when those cells, multiples of them, come together in early development of an embryo and what wonders can they do. When the blastula cell mass during early development, as epithelial sheets, morph into different shapes and structures, we call this stage "**gastrulation**". Gastrulation is defined by the collective cell movements that mediate formation of developmental architecture, importantly the three germ layers.

Cases of Gastrulation: Fly-fishing

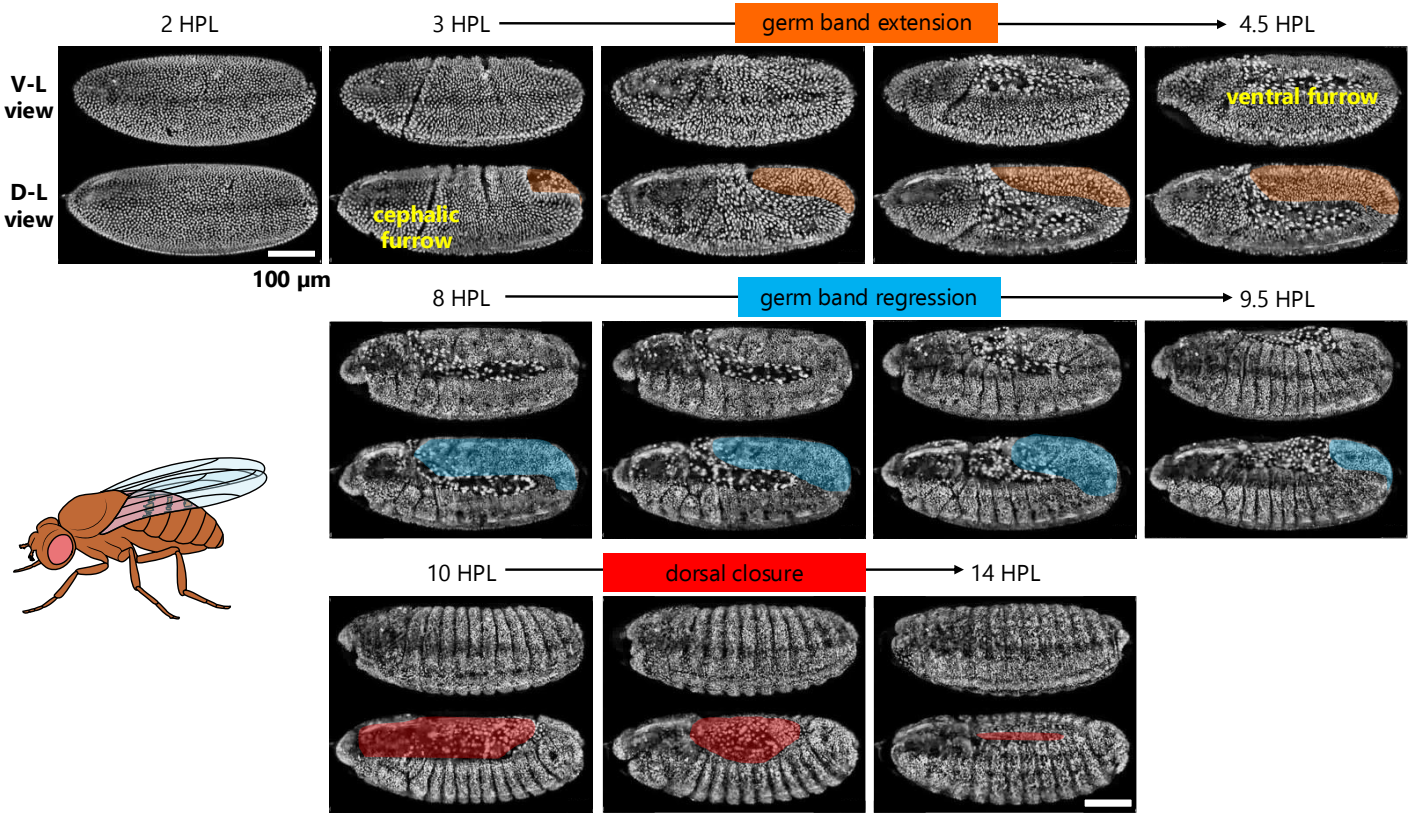

Figure 7.1 | *Drosophila melanogaster* gastrulation as observed with fluorescent cell nuclei marker (V-L: ventrolateral, D-L: dorsolateral views). HPL: hours post egg-laying. Scale bar is 100 microns. Timelapse snapshots are reproduced from light-sheet microscopy in P. Keller Lab, Janelia Research, MD, USA.

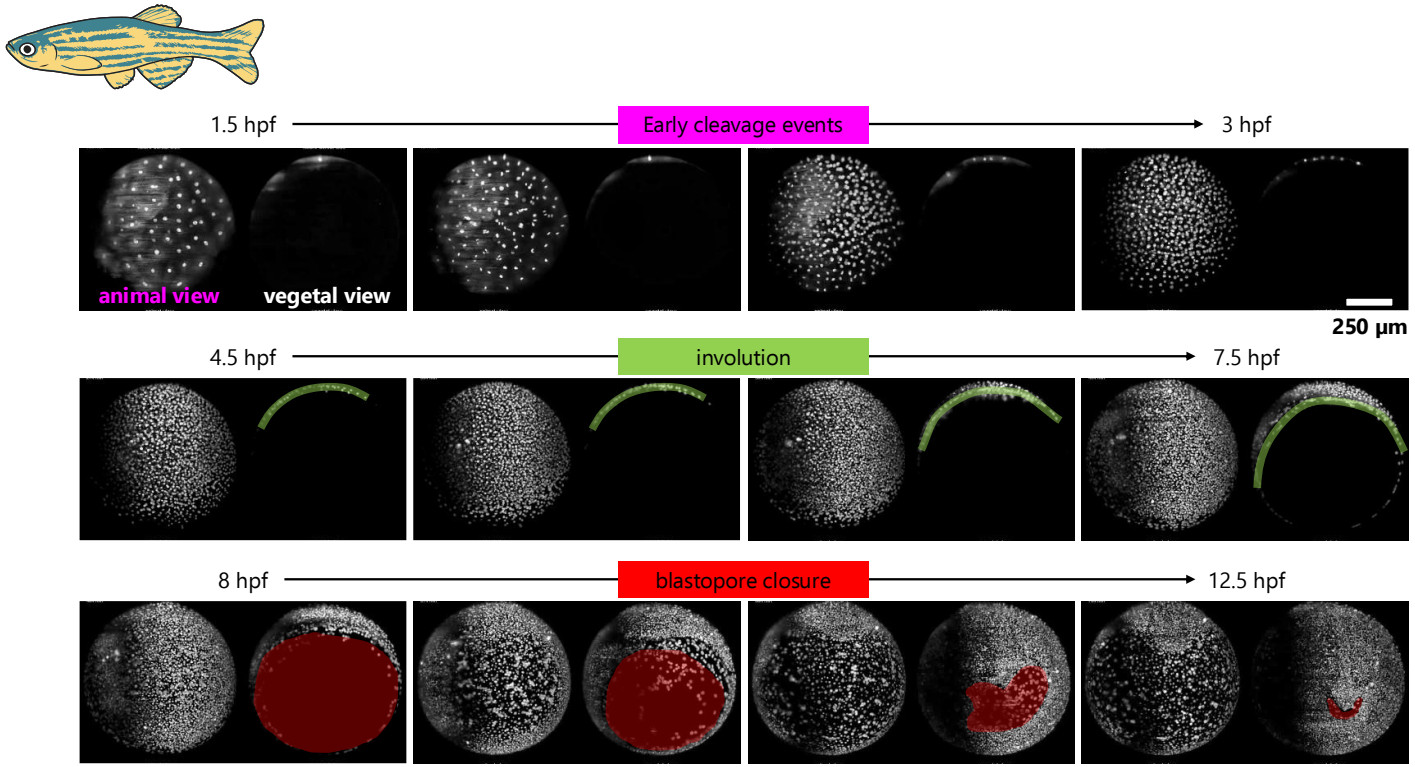

Figure 7.2 | *Danio rerio* gastrulation as observed with fluorescent cell nuclei marker (animal and vegetal views). hpf: hours post fertilization. Scale bar is 250 microns. Timelapse snapshots are reproduced from light-sheet microscopy in P. Keller Lab, Janelia Research, MD, USA.

*Careful observation of how gastrulation unfolds for two organisms will be useful to get a grip on underlying principles. First example is the fruit fly, Drosophila melanogaster (**Figure***

__7.1__). In fruit flies, gastrulation is marked with a rapid expansion of the epithelial sheet referred to as __germ band extension__. Afterwards, the sheet comes back as more defined segments, which is called __germ band regression__. The opening at the regression is going to zipper, which is called __the dorsal closure__ since it is at the dorsal surface of the embryo. How do the germ layers form? Following extension, the cells at the "bottom" of the embryo dive "inside" the __ventral furrow__. They further dive through the __cephalic furrow__, which separates the head tissue from the abdomen. Those invaginations of the epithelial sheet results in the formation of the mesoderm, the mesoderm __invagination__. The tissue that forms the endoderm similarly gets invaginated but from the very end of germ band extension, and afterwards the germ band regresses. Broadly, the quick expansion of the epithelial sheet causes invagination of material that will become the __endoderm__ and invagination of the material that will form the __mesoderm__, and regression of the tissue ultimately drives the segmentation of the outer germ layer, the __ectoderm__. The expansion, invagination, and regression movements of these cell bundles are the collective phenomenon of gastrulation.

Next, we can look into zebrafish, Danio rerio, gastrulation as an alternative model with a well-defined animal-vegetal pole (__Figure 7.2__). In the beginning, we see that the animal pole is enriched for nuclei, while the hemisphere of the vegetal pole harbours none. During development, the cells at the animal pole will expand into a single-layer epithelial sheet whose division will force their __expansion__ downwards towards the vegetal pole. The hole of the blastula, the __blastopore__, closes and then we see the cell migration events that accumulate cells in the head-to-tail embryo __axis__. Effectively, as the blastopore lip closes, a certain layer of cells goes inside, and we call that __involution__, which is when the mesoderm and the endoderm are going under the ectoderm. The cell __ingression__ and __intercalations__

along the head-to-tail axis are what further drive the growth of the embryo axis, i.e., ***axis extension****.*

These two examples should articulate that, between different species, events happening during the blastula stages can be vastly different. But, as gastrulation continues, we will see that the morphology of the different embryos will be similar to each other – i.e., patterned into three embryonic germ layers. This will particularly put various species for vertebrates under very similar architecture through gastrulation.

Table 7.1. Ways to reshape epithelial tissues

Descriptive Terms	Biological Processes

• Convergent extension
• Epiboly
• Invagination, involution, ingression
• Closure
	• Directed cell division
	• Cell intercalation
	  ○ Radial intercalation
	  ○ Planar intercalation
	• Collective cell migration
	• Boundary compartmentalization
	• Adhesive sorting
	• Fluid-solid (jamming) transition

How are the epithelial sheets reshaped during gastrulation? The table above outlines the primary biological processes and the forms of epithelial reshaping observed with descriptive terms used. We will now dig deeper into the underlying biological processes until the end of Chapter 8.

1. Cell intercalations

Cell intercalation events result in the loss of dimensionality between two distinct layers of epithelia. How the two epithelial sheets merge and the directionality of their subsequent division events determine the type of intercalation that takes place.

Radial intercalation: This occurs when two epithelial sheets, which are stacked on top of one another for example, collapse into each other and form a single 2D sheet that grows in both perpendicular axes. Radial intercalations result in the thinning and **spreading** of the epithelial sheet.

Example: The **epiboly** of zebrafish to expand over the yolk.

Planar intercalation: This occurs when two sheets, which are side-by-side for example, collapse into one another to form a tissue that only expands in one axis. The tissue on one axis gets shorter from the intercalation event, while the perpendicular axis continues to elongate. This is also referred to as **convergent extension**.

Example: The germ band extension during fruit fly gastrulation.

Biophysics of extension processes: Intercalations are not static events; cells in contact can lose that adherence and find new binding partners. This mutability allows for the rearrangement of neighbouring cells without needing to change the composition of the epithelial sheet. To better understand this concept, we will focus on the three major components that allow for germ band extension to take place.

Let us consider the simplest example where we have four neighbouring cells: two green cells that do not contact each other and two grey cells that are adhered to one another. If the two grey cells lose their connection and the green cells adhere to one another, planar intercalation takes place with one axis thinning and the other extending. When you have four cells, two which lose a connection and the other two that form a connection, a **T1 transition** takes place. This coupling of a junction collapse with the formation of a new junction allows for "neighbour exchanges" without dedicated cell rearrangement events. The first component of germ band extension is those T1 transitions.

However, a junction can collapse between more than four cells, and when this happens all the cells share a junction in the centre. This "all-cell" junction that is transiently formed is called a rosette, and it is a mechanically unstable structure. This instability is a consequence of physics – modelling conducted on soap bubbles revealed that the stable coordination number of a vertex, in a two-dimensional space such a sheet, is three. That is, a single vertex is most stable when it is in contact with three distinct edges. The findings that

were taken from soap bubbles hold true for cells. For rosettes, the cells will reshape and form an elongated sheet by migrating to a perpendicular axis. This **rosette resolution** results in the tissue flipping the orientation of cells, and the exchange of neighbouring cells. Rosette resolution is the second component that helps drive germ band expansion.

Both T1 transition and rosette resolution events are driven by the mechanical instability of having more than three contact points following a junction collapse. A large takeaway is that the physical properties of the membranes drive the transition events which lead to the reorientation of epithelial cells.

The third component, after T1 transitions and rosette resolutions, which drives germ cell expansion is **cell stretching** and global tension forces. For example, as the midgut invaginates from the posterior end of the fruit fly, the force of the movement expands the epithelial sheet itself.

We are now going to put our magnifying glass on the first two components further: T1 transitions and rosette resolutions. What biological features control the collapse of junctions and their resolution aligned along the full embryo axis - is it genetics or biophysical properties? The collapse of junctions is primarily controlled by actin-myosin stretching. As the actin filaments stretch, they enforce a stretching of the cell that drives junction collapse. So, we can consider junction collapse as a consequence of the movement and structural components of each cell.

What about the resolution of these junction collapses? An important feature in the resolution of these events is the orientation of adherence. For germ layer extension, we need the transitions to be oriented along a single axis. This is where biology comes back in. When we look at the shape of tissues, we can define certain interfaces between the cells as horizontal or vertical interfaces. This adds an additional layer of polarity; actomyosins are heavily oriented along the vertical interfaces, while the horizonal interface is enriched for cadherins. This polarity influences how junctions form. The horizontal lateral membranes, rich with adherens junctions, are where cell-to-cell contact sites form and the cell is shrinking along the other axis because of myosin stretching on the other lateral membrane. Basically, epithelial cells now exhibit polarity both from a top-bottom (apical-basal) layer and side-to-side layer which we refer to as lateral polarity. This lateral polarity helps determine the axis by which these resolutions take place following a junction collapse.

What maintains **the lateral polarity** of a cell? Considering that cadherins are membrane bound, what is limiting their diffusion throughout the membrane? This is driven by both intercellular positive interactions and intracellular negative interactions. For intercellular positive interactions, we know that membrane bound proteins want to interact with each other. An unbound protein is less energetically favourable than in its bound state. If cadherins begin to accumulate on a membrane, we would expect that cadherins will accumulate on the neighbouring cell's membrane so that the proteins can bind. This is the basis of intercellular positive interactions, that as one surface gets polarized, the neighbouring surfaces of the other cells will also be driven towards same polarization. The same goes for actomyosin – as actomyosin get enriched to one side of the membrane, the physical force will increase not only on that given surface, but the surfaces it comes into contact with as well (following Newton's third law). Conversely, the accumulation of cadherins on a membrane would decrease the proportion of myosins there, and vice versa. This intracellular negative interaction is very similar to lateral inhibition, where if you have one pool of molecules "winning" it will lower the proportion of the others, causing them to "lose", eventually causing separation and polarization.

The combination of **positive intercellular** and **negative intracellular interactions** can mediate lateral polarization between neighbouring cells, but to maintain this polarity throughout the epithelium, an additional external cue is needed. This external cue can arise from other neighbouring cells in the form of diffusive signalling, it can be a morphogen gradient that biases the tissue along one axis, or it can be the stripy expression of some genes, which is the case for the fruit fly germ layer extension.

2. Boundary compartmentalization

We understand that the horizontal membrane, rich with cadherin proteins, drives the formation of junctions between epithelial cells. The actomyosin structures that run perpendicular to this axis and define the vertical lateral membrane play a key role in mediating tissue separation. The polarity between the vertical and horizontal interfaces of epithelial cells gives rise to actomyosin tension, which drives tissue separation and subsequent **boundary compartmentalization** or compartment boundary formation. Recall the germ band of the fruit fly, which begins as an unstructured cellular mass, but following regression, forms defined segments. Once these are formed, the long actomyosin cables going through the tissue deter cells of different

compartments from mixing with each other. This way, different boundaries are defined in the developing embryo due to the tension derived from the polarity between interfaces.

The planar cell polarity system

The concepts of structural polarity within epithelial cells discussed so far were primarily focused on the example of fruit fly germ band extension. Now is a good time to expand on a more broadly relevant concept: the planar cell polarity, as a conserved process across animal phyla. To appreciate the fashioning through planar cell polarity system, we will look to the hairs found on the wings of fruit flies, the **trichomes**. In wild-type fruit flies, the hairs on the wings run in parallel with one another, but in a *vangogh* (*vang*) mutant the trichomes instead exhibits a distinct swirling phenotype. This turbulent hair organization, which resembles the sky from Van Gogh's 'Starry Night', is caused by the absence of a functional Vang product. The parallel organization of wild-type fruit fly hairs is the consequence of planar polarity, and in this process, Vang is an important player.

How does the planar cell polarity system work? Let us begin by looking at the membrane composition of a single cell (**Figure 7.3 inlet**). In this cell, we have side A (**right side**), which has two transmembrane proteins in complex with one another: the apical cadherin protein flamingo (Fmi) and the Vang receptor. On side B (**left side**), we have another transmembrane complex, this time, between another Fmi molecule and the Frizzled (Fz) receptor. The Fmi-Fz complex of this cell can form a connection with the Fmi-Vang complex of another cell through homophilic interactions between Fmi molecules. Once an intercellular Fz-Fmi-Fmi-Vang complex forms, additional members are recruited to the cytoplasmic interface of the single cell we are looking at. On side A, where the Fmi-Vang complex is interacting with another cell's Fmi-Fz complex, the Prickle protein (Pk) binds to the cytoplasmic interface of Vang. On side B, where the Fmi-Fz complex is interacting with a neighbouring cells Vang-Fmi complex, the disheveled protein (Dsh) binds to the cytoplasmic interface of Fz. Now, within the single cell, side A has a Fmi-Vang-Pk complex, and side B has a Fmi-Fz-Dsh complex. The recruitment of Dsh allows for this protein to activate actin polymerization on side B of the cell. However, the complex of side A interferes with this activity; specifically, the Fmi-Vang-Pk complex inhibits Dsh activity. It is only through the recruitment of an additional protein factor, Diego, the side B complex can override this Vang-Pk inhibition. Diego, in this Fmi-Fz-Dsh-Diego complex, represses the inhibitory activity of Vang-Pk. This

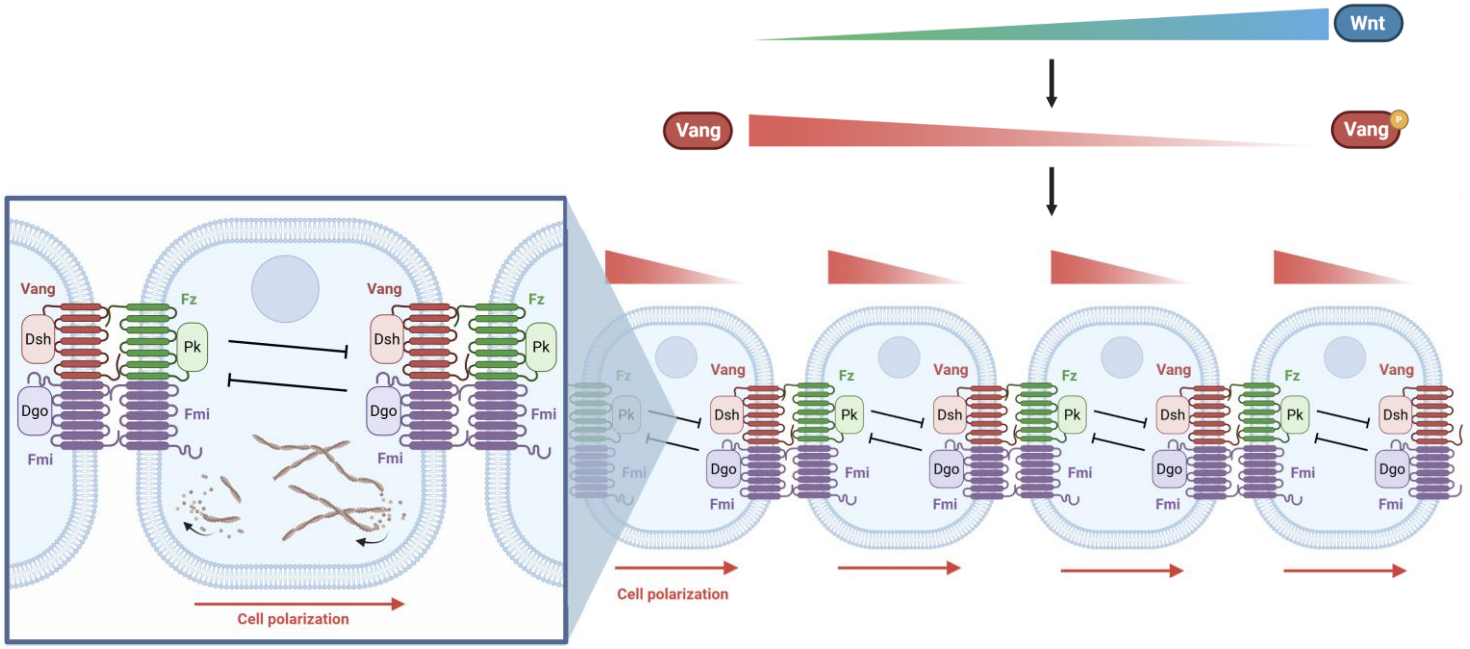

Figure 7.3 | Planar cell polarity (PCP) system.

double negative inhibition activates Dsh and instigates actin polymerization. This orients **actin polymerization** to a single "side" (**Figure 7.3 inlet**) or lateral surface of the epithelial cell and helps in establishing planar polarity. The membrane complexes contribute to intercellular connections, but the inhibition pathways occurring within a single cell localize these complexes on opposing membrane surfaces. **Intercellular positive reinforcement** and **intracellular negative regulation** hand-in-hand polarizes the epithelium across the cells.

1. PCP system across the tissue

Finally, the PCP system is controlled by a global cue to work across the full-scale tissue: the Wnt signalling pathway. Wnt signal results in the phosphorylation of Vang into $Vang^{P}$; $Vang^{P}$ cannot inhibit Dsh activity (**Figure 7.3**). Let us consider the consequences, again, in the context of a single cell comparing between side A and side B. Let us assume that side B is closer to the source of the Wnt signal. This side will accumulate more inactive $Vang^{P}$ than side A; establishing a Vang activity gradient between opposing "sides" of a cell, *i.e.,* within each cell, that is opposite to the gradient of the Wnt signal. Each cell in the tissue will experience the "opposite" gradient of Vang activity in accordance with the relative Wnt levels. Globally, this results in a saw-tooth pattern of Vang activity throughout the tissue that is polarized to the same side of different cells within an epithelial sheet. The global cue enables the **synchronized orientation** of planar cell polarity and allows for all the hairs in the wing of the fruit fly to grow parallel to one another.

We should now come to understand that planar polarity allows for cells coordinate their inner cell orientation towards a plane and harmonize their arrangement. It is through this unison that we have the adhesive coupling we see in epithelial sheets, convergent extension we see in very early development, and the rosette resolutions as well. As cells are able to polarize themselves, multiple environmental cues allow them to share an orientation and coordinate their polar structures. This resembles how electric dipoles with coupled negative and positive charges will align themselves according to their neighbours by opposite charges attracting each other. An external electric field can then align all the polarity of the dipoles and polarize the material. Planar cell polarity (PCP) is a deeply conserved system and involved in many steps of embryological development.

2. PCP system and neural tube formation

During vertebrate development, PCP is required for the neural sheet to fold into neural tube. As you may recall, outer cells of the embryo, the ectoderm, give rise to both the epidermis and the neural tissue. Neural ectoderm is initially a plane of epithelial cells which we call neural sheet. Extending the sheet in the appropriate axis, the epithelial layer will begin to fold in on itself and form the neural plate which folds into the neural tube.

Cells are brought to the midline, mediolaterally shrinking the sheet and extending the tissue over the anterior-posterior axis: **convergent extension**. The neural tube can then differentiate into diverse types of neurons giving rise to the central nervous system, the spinal cord, *etc*. At that initial phase of development, the epithelial sheet needs to undergo convergent extension and elongate along the body axis. Researchers tracked those tissue movements in mice embryos over the course of 8 hours; for wild-type embryos as well as *Van gogh like* (*Vangl*) and *Ptk7* mutants. Ptk7 is a signalling partner for the Frizzled receptor. While the tissue converged in one axis and extended in the other along the A-P direction in the wild-type, in mutants with deficient PCP system the convergent extension of neural sheet did not occur.

The next step, which requires the folding and invagination of the sheet is also driven by PCP and is required for the closure of the neural tube. What is the role of PCP in folding? Key molecule here is Celsr1, which is the vertebrate homologue of Flamingo (Fmi) that accumulates at the adhesion points between the epithelia. Cells apically localize the Celsr1 during neural tube invagination. Having the Celsr1 polarized, the phosphorylation of myosin is activated in the apical actin-filament rich cell surface. The phosphorylation results in the constriction of each epithelial cell at the apical surface, transitioning the cell into a "squeezed" **wedge-like shape**, where the basal surface is wider than the apical. This collective "wedging" of the apical surface of the sheet causes the tissue to invaginate.

Using siRNA, a knockdown of Celsr1 at this developmental stage revealed a loss of closure and abnormal spine development. In fact, one of the most prevalent PCP-related birth defects, spina bifida, results from the neural tube not being properly closed and zippered. Different forms of **spina bifida**, from spina bifida occulta to myelomeningocele, are linked to different mutations in the Vang receptor. Those mutations broadly result in the incomplete invagination of the neural sheet, leaving a gap within the spine during the 4th week of pregnancy in humans. It is striking that the same protein complexes that orient the hairs of the fruit fly play such an important developmental role for humans as well, using same set of design principles.

FURTHER READING

Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. P. J. Keller *et al.*, **Science**, 2008, DOI: 10.1126/science.1162493.

Cellular, molecular, and biophysical control of epithelial cell intercalation. A. C. Paré, J. A. Zallen, **Current Topics in Developmental Biology**, 2019, DOI: 10.1016/bs.ctdb.2019.11.014.

Wnt-Frizzled/planar cell polarity signalling: cellular orientation by facing the wind (Wnt). Y. Yang, M. Mlodzik, **Annual Review of Cell and Developmental Biology**, 2015. DOI: 10.1146/annurev-cellbio-100814-125315.

Chapter 8: GASTRULATION MOVEMENTS

"Caminante, son tus huellas
El camino, y nada más;
Caminante, no hay camino:
Se hace camino al andar.
Al Andar se hace el camino,
Y al volver la vista atrás
Se ve la senda que nunca
Se ha de volver a pisar.
Caminante, no hay camino,
Sino estelas en la mar..."

– Antonio Machado (1875 – 1939)

Our next journey aims to understand how gastrulation shapes embryos into a patterned structure. We will begin with revisiting the cell and tissue movements. All embryos before gastrulation can be thought of some unstructured slabs of epithelial tissue in 3-D space. Some will be more planar, *e.g.,* bird, reptile, or human embryos, some will be engulfing a yolk in ellipsoidal shape, *e.g.,* holometabolous insects like fruit flies, some will sit like a dome-shape on top their yolks, *e.g.,* many fish embryos. Those tissue slabs will be molded with 1- epiboly, 2- emboly, and 3- convergent extension movements during gastrulation.

1. Epiboly

Epiboly is expansion of the tissue in two dimensions while thinning in depth. This is mainly driven by **radial intercalation** of the cells. **Directed migration** of leading-edge cells towards vegetal pole helps the process. A classical example discussed earlier was zebrafish gastrulation during which animal pole cells engulf the yolk fully by completing the epiboly stages.

2. Emboly

Emboly is when a sheet of tissue begins folding and undergoing itself, forming two layers. Emboly is driven by ingression, invagination, or involution of the cells. When you have individual cells not sticking together, going down, or dripping from the sheet that is called an **ingression**. This happens when one has

extensive tension on the tissue *e.g.,* chick for which the embryo develops on a stretched plane under vitelline membrane. That stretched plane causes the epiblast, the outer layer cells, to go inside one by one; like the cells are dripping from the groove called primitive streak on the hypoblast. Ingressing epiblast cells eventually give rise to all three germ layers: the ectoderm, endoderm, and mesoderm. When the whole tissue is bending and going down together, with the apical constriction of cells, we call this **invagination**. In fruit flies we spoke about invagination of the gastrulating tissue, that the prospective mesoderm was invaginating at the ventral furrow and the cephalic furrow. Another example is neural tube invagination we discussed recently. Lastly, **involution** is inward rolling of the cells which, more or less, is a mixture of ingression and invagination. In fish and amphibian embryos, cells at the yolk margin or blastopore involute under the outer layer of cells, giving rise to the mesendoderm inside, and ectoderm outside.

3. Convergent extension

Thirdly, planar intercalation, also known as mediolateral intercalation, depending on the positioning of the axis, drives convergent extension. We discussed the **T1 transitions** can cause tissues to expand. We also know that, during gastrulation, cells coagulate into the midline of the embryo. Besides the convergence and extension, this **directed migration** brings cells to the midline to give rise to a structured elongating body axis.

Movements of gastrulation eventually puts embryos of different vertebrate species into **a similar body plan** regardless of different egg shapes they begin from. Embryonic midline is marked by primitive streak/notochord. A longitudinal body axis is established with prechordal plate in the anterior (head) side. Mesoderm flanks bilaterally to the midline and more lateral regions of the germ ring enter underneath to specify endoderm. It is important to keep tissues structured as those collective movements of gastrulation mold the embryo into shape. We will discuss biochemical and biophysical mechanisms driving and maintaining tissue patterning next.

Adhesive sorting

In 1955, Townes and Holtfreter took the early embryo of a frog, and compared the organization of germ layers, most specifically the neural ectoderm and the endoderm, between an unaffected embryo and one whose germ layers have been scrambled. In the normal organization of the frog embryo, the neural fold

ectoderm sits atop the endoderm. With apical constriction, invagination, and neural tube folding, the radial body plan follows the inner endoderm being maintained, a middle layer of neural cells including the tube and mesenchymal neural crest cells, and an outmost ectoderm layer that later forms the skin. Surprisingly, when the neural fold ectoderm and endoderm were scrambled, the ectoderm cells would form individual islands that coalesced into a single larger one eventually. This large neural ectoderm island ended up organized very similar to the unperturbed embryo. Their experiments revealed that the cells belonging to different germ layers can self-organize and introduced the concept of 'Gewebeaffinität' (**tissue affinities**), which suggests that ectoderm cells have a greater affinity towards other ectoderm cells, endoderm cells with endoderm cells, and so on. Much later, it was revealed that this affinity is greatly influenced by the surface tension of the tissue; a higher surface tension means that the structure will be harder to sandwich or squeeze, while a lower surface tension means it will be much easier to squish. In the 1990s, scientists aimed to compare the surface tension of different tissues, *e.g.,* retina, liver, heart, pigmented epithelium, and limb bud. Mixing cells from those tissues in pairs revealed cell sorting in action; depending on surface tension differences, softer tissue cells sit outside engulfing the stiffer tissue cells. Later investigations showed this variability was due to the differences in **cadherin** levels rather than the types of cadherins expressed in the different cell types.

What is the impact of variable cadherin levels, and how does this relate back to the tissue affinities? An experiment was conducted where two cell types were suspended in a homogenous mixture. The first group of cells expressing a red-fluorescent marker had very little cadherin expression, while the second group of cells expressing a green marker were induced to expressed double the cadherin levels of the first group. After 24-hours, it was observed that the cadherin-rich green cells congregated to the middle of the cell mixture, while the cadherin poor cells surrounded the middle structure. This phenomenon is known as '**adhesive cell sorting**'. Why this sorting takes place in accordance with cadherin levels is rooted in physics: the minimal energy principle. This principle states that, due to energy-time symmetry, energy tends to be minimized if you let a system evolve or dynamically change over time.

To illustrate this, let us focus on the cell population that can be grouped into two categories: high cadherin (HC) and low cadherin (LC) (**Figure 8.1**). The cadherins are able to diffuse around the membrane, such that they would become evenly distributed. This population begins as a mixture, where the cadherins of

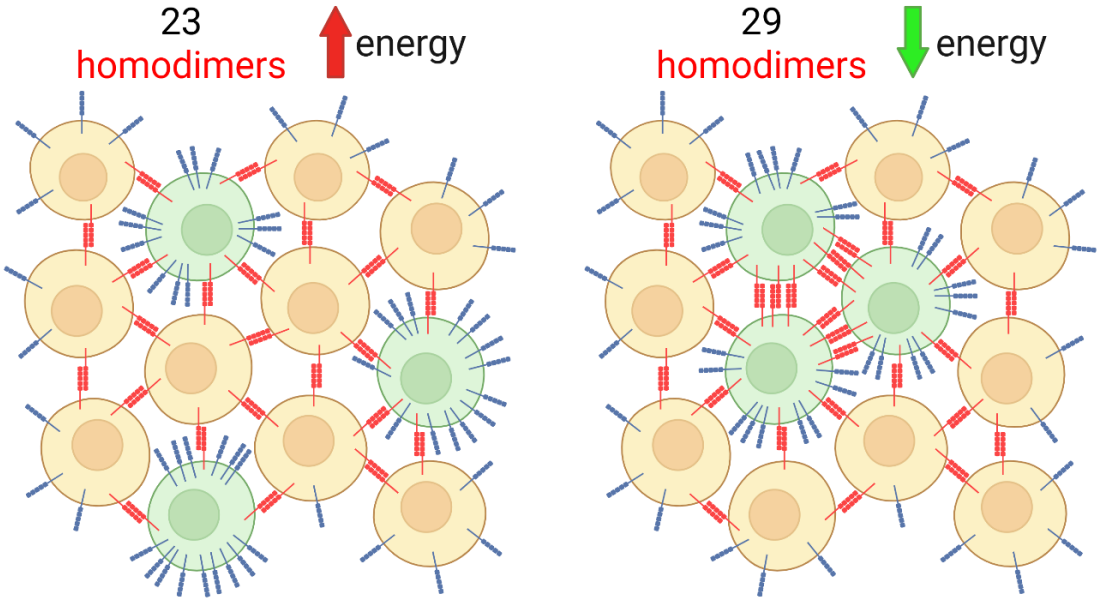

Figure 8.1 | Adhesive cell sorting due to differential cadherin expression between green (HC) and orange (LC) cells. Transmembrane cadherin receptors (blue) forming homodimers are highlighted in red. Arrangement on the right side is energetically favourable.

the HC cells interact with the cadherins of the LC cells. If the HC cells are neighboured only by LC cells, there will be cadherins in the HC membrane that are left unbound, which is energetically unfavourable. It is more stable for a cadherin protein to interact with another cadherin protein, than for it to remain unbound. If we allow this system to evolve over time, by the minimal energy principle, we will observe that the population of cells lowers the energetic cost of its arrangement by having HC-HC cell interactions, while the LC cells occupy interactions among each other. HC cells with higher surface tension caused by higher cadherin levels will congregate to the centre of the population, which will maximize cadherin dimerization events, and lower the energetic cost of the arrangement.

Cadherin's role in epiboly

What are the consequences when cells are unable to appropriately sort? E-cadherin mutant zebrafish are shown to be unable to complete gastrulation. Embryos homozygous recessive for this allele, called *half-baked*,

are able to begin the gastrulation and spread over the yolk but are unable to complete this process and get stuck partway through. The loss of the E-cadherin molecule interferes with the radial intercalation and also the collective cell migration required in epiboly.

In the large epithelial slab of tissue before gastrulation, the cells in the outer most layer highly express E-cadherin while that underneath (radially closer to the yolk) express less. However, we know that it is energetically unfavourable for these HC cells to be at the surface of a cellular mass, and so they try to find each other. As the LC cells are pushed from the centre outwards, their cadherin levels also begin to increase when they move towards the outer layer. This combination of **cadherin gradient** and adhesive sorting results in a unique outcome: the circular movement of HC cells finding each other on one axis the LC cells coming upwards and increasing in cadherin levels, creates positive feedback that results in tissue spreading. Ultimately, this spreading will go over the entirety of the yolk, closing the blastopore and establishing three germ layers.

Biophysical mechanisms of vertebrate axis elongation

All vertebrate embryos require axis elongation to take place, from fish to frogs, chicks to humans. Posterior embryonic tissue undergoes a physical state transition during the collective migration movements of gastrulation. This **fluid-to-solid jamming transition** is a recurring theme we observe in natural phenomena: the stalactites in caves, the icicles elongating from rooftops in the winter, and even in traffic jams share a similar underlying cause for elongation. A structural continuum, where one side is entirely solid while the other is entirely fluid, is present in all those very distinct cases. As the highly motile fluid phase hits into the solid phase, this causes the fluid to undergo jamming transition. Consequently, the structures expand along one axis unidirectionally. The traffic jam extends further into the highway, the icicle gets longer and longer, so on and so forth. Similarly, we see this happening in embryos.

Measuring the physical features of the anterior-posterior (head-to-tail) axis in zebrafish embryos, researchers realized that at the tail bud there is greater extracellular space between the cells allowing the cells to be quite motile. As you move along the axis towards the head, the space between cells decreases and they become increasingly more packed and less motile. Like the dripping stalactites in caves grow, the tail bud cells are initially quite fluid, but the jamming transition causes them to solidify, losing space and motility over time.

The embryonic body axis expands unidirectionally towards posterior. This is not isolated to axis elongation in zebrafish; the jamming transition is also observed in chick embryo development. When researchers isolated the chick presomitic mesoderm, the unsegmented mesoderm tissue in the tail, and confined it in between two plates, they have realized the tissue elongated unidirectionally towards posterior. Similarly, cells had more extracellular space in the posterior and displayed higher motility. Expanded intracellular space and higher motilities create higher mechanical stress driving the elongation of the body axis.

What does bias the posterior tail bud cells to be more motile and result in larger intracellular space? The culprit is fibroblast growth factor (FGF) signalling. FGF signal is remarkably high in the tail, but much lower towards the segmented mesoderm in the anterior. **FGF signalling** increases cell motility, creating the motility gradient. FGF signal further expands extracellular space in between cells by triggering the production and secretion of hyaluronic acids. As the FGF signal decreases towards anterior, hyaluronic acid synthesis is decreased, and the cell-to-cell pressure decreases allowing cells to become more tightly packed.

Our investigation of the biological (biochemical and biophysical) processes reshaping epithelial tissues during the gastrulation of an embryo, which were summarized in **Table 7.1**, ends here. Concluding the ways gastrulation reshapes the embryo, we will move ahead with how 3-D coordinates and patterns emerge in development in the next chapter.

FURTHER READING

Directed movements and selective adhesion of embryonic amphibian cells. P. L. Townes, J. Holtfreter, **Journal of Experimental Zoology**, 1955, DOI: 10.1002/jez.1401280105.

Conserved patterns of cell movements during vertebrate gastrulation. L. Solnica-Krezel, **Current Biology**, 2005, DOI: 10.1016/j.cub.2005.03.016.

A fluid-to-solid jamming transition underlies vertebrate body axis elongation. A. Mongera *et al.*, **Nature**, 2018. DOI: 10.1038/s41586-018-0479-2.

Chapter 9: EMBRYO'S CARTESIAN COORDINATES

"IT SEEMS THAT NATURE IS CONFINED WITHIN CERTAIN LIMITS AND HAS FORMED LIVING BEINGS WITH ONLY ONE SINGLE PLAN, ESSENTIALLY THE SAME IN PRINCIPLE, BUT THAT SHE HAS PRODUCED VARIATION IN A THOUSAND WAYS IN ALL HER ACCESSORY PARTS..."

– Étienne G. S.-Hilaire, *in* Magasin encyclopèdique, 1796

We have so far observed the gastrulation from a morphogenetic perspective related to the cell and tissue mechanics, form, and movements. We will now move on to a central topic in developmental biology: axis specification. What is an axis? How do we define it for an embryo?

Once an embryo gastrulates it has greater structure and shape, with its three orderly germ layers, compared to the blastula which was closer to a blob of cellular mass. At that point, we can clearly speak of Cartesian coordinates with three perpendicular axes. René Descartes first described the Cartesian coordinates in analytical geometry. The basic idea is if someone wants to describe the motion and change of form for an object, they need to set the scene first: *i.e.*, three perpendicular axes: x, y, and z, for the three-dimensional space the matter occupies. For an embryo, those are the anterior-posterior (AP, head to tail), dorsal-ventral (DV, top to bottom or back to front), and the medial-lateral (ML, right-to-left) axes.

Note that this is something we impose on the embryo to describe it. The central question of this chapter is whether the cells really care about any of them? Are those Cartesian coordinates, as I look at the geometry and define for the embryo with my human brain, really existing for a cell living in that embryo and patterning it?

To be able to speak of presence of axial coordinates, three conditions must be met. The first is for the form to have **symmetry breaking**. If the form were perfectly symmetrical, an axis could not be defined with positive and negative sides, *e.g.*, one side being anterior *vs.* the other being posterior. The second feature is for the coordinates within the plane to have **positional information**, such that the location of a point along the axis means something beyond the mere asymmetry. The final feature is that there is a structural

architecture associated with each axis; that there is a **body plan**, for instance when one compares the dorso-ventral axis of embryo 1 with the dorso-ventral axis of embryo 2.

"Unity of the body plans" hypothesis

We will begin with the concept of a body plan. This idea was first clearly formulated by the French scientist Étienne Geoffroy Saint-Hilaire in 1820s. He proposed that nature presents one plan of construction for animals which is same in principle but may be varied in accessory. Those ideas are famously recorded in his debates with French biologist Cuvier in Paris. During the debate, he would dissect an arthropod species, the crayfish, to make his point that the back-to-front axis of arthropods are patterned similar, but at the flipped order, with human back-to-front axis in terms of the placement of inner organs. If we were to compare now, the dorsal-ventral body plan of a zebrafish, a vertebrate, and a crayfish, an arthropod, we would see that Saint-Hilaire was correct in his thinking; the body plans for the two dissimilar species were very conserved, only that the DV axis is flipped (**Figure 9.1**). The central nervous system (CNS), which goes through the ventral line in arthropods, matches the dorsal side observed in vertebrates. Comparing the position of the CNS and the mouth for both organisms, we see that the orientation is opposite. Inner organs come in flipped order, implying a conserved general body plan.

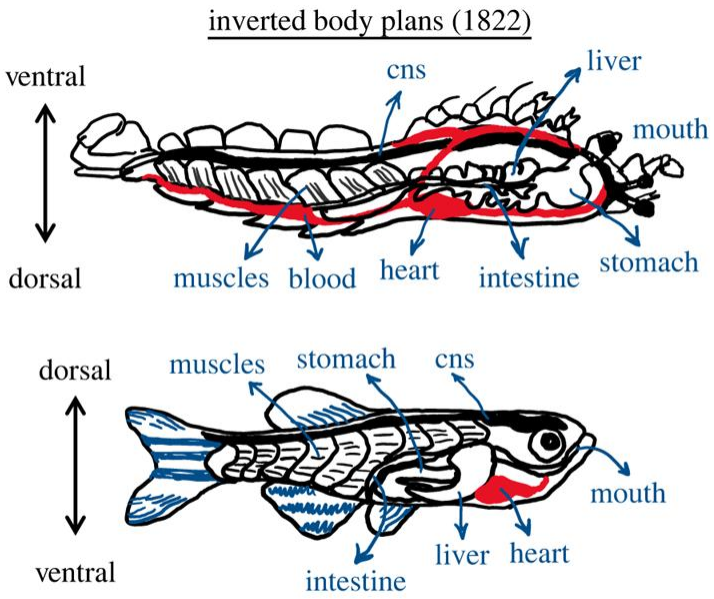

Figure 9.1 | Inverted body plans model between invertebrates (crayfish) and vertebrates (zebrafish as example). An upside-down crayfish matches in order of its internal organs with zebrafish. Adapted from Figure 1a, in Simsek and Özbudak, Open Biology, 2022 (DOI: 10.1098/rsob.220224).

However, for almost two centuries, the arguments for an "embryonic body plan" were restricted to anatomical evidence.

How about the symmetry breaking? This happens very early on in development, usually when the sperm enters the egg. We have previously covered other features such as the maternally deposited factors that lead to polarization or asymmetry upon egg activation. However, an additional mechanism shall be present to ensure division axis of fertilized zygote gets lined up with those maternal factors. This is the cortical rotation. Using amphibian fertilization as an example, when the sperm contacts the egg, we observe a ~30° cytoplasmic rotational flow. While this angle is amphibian specific, rotational cortical flow is a general concept from amphibians to fish and mammals and is essential to establish first embryonic axes. That rotational flow brings certain maternally deposited factors, *e.g.*, Dishevelled as discussed for sea urchin embryos in Chapter 5, opposite to the sperm entry site. A unique feature of amphibian eggs is that its animal side is heavily pigmented, and the vegetal side lacks much of pigmentation. This makes tracking the cortical rotation upon the sperm entry easy. When the sperm activates the egg and the cortical rotation takes place, the 30° cytoplasmic rotational flow drives the formation of the grey crescent zone, *i.e.*, mixed pigmentation, in between the animal and vegetal hemispheres.

The organizer

That brings us to what one could say the most impactful experiments ever done in developmental biology: Hans Spemann and his lab's works on newt development in the early 1900s. In 1900, Spemann would try to interfere with and alter the development of the newt by constricting the early embryo with thin lassos made from his baby daughter's hair. When Spemann performed the constriction at 8 to 16-cell stage, between the animal and the vegetal poles, this resulted in one or few nuclei of the animal pole during the fourth division escaping into the vegetal region. After these nuclei escaped to the vegetal pole, he observed that rather than one, a twin of newts would develop. Following this, he focused on the grey crescent and saw that when he conducted the experiment in the earlier 2 to 4-cell stages, it was only when the grey crescent was present in both halves could he see two newts forming. If the grey crescent were restrained to one side of the constriction

event, only a single newt would arise alongside an additional formless mass comprised of some blood and ectoderm absent of any organismal structure. Spemann referred to this a 'belly piece'.

Spemann's constriction experiments in 1900 suggested that the grey crescent was important in embryological development. We now know that many maternally deposited proteins and RNA, including the transcripts important for Wnt signalling, localize to the grey crescent following cortical rotation. Protein factors, such as the GSK-3β binding protein (GBP) and Dsh, are carried to the grey crescent by the microtubule molecular motors, kinesins. Wnt ligand RNA itself, on the other hand, diffuses in that direction following cortical rotation. The presence of the Wnt ligand RNA, Dsh, and GBP at the grey crescent results in this site being really high in Wnt signalling, with elevated nuclear β-Catenin levels in the presumptive dorsal zone next to the grey crescent. These together are required for and drive embryo formation.

Hilde Mangold-Pröscholdt here comes into scene. A student of Spemann's who tragically passed away two years after completing her Ph.D., she conducted a series of (two-hundred and fifty-nine recorded in the lab notebook) transplantations that illuminated the role of the grey crescent in embryological development. Hilde Mangold-Pröscholdt used two newt species, each with their unique egg pigmentation for those experiments, enabling tissue tracking post-transplantation. Knowing the grey crescent is the future side of the blastopore lip where the mesendoderm is invaginating during gastrulation, she transplanted that site from one embryo onto the opposing side for the epidermis on the second embryo, essentially forming an embryo with two distinct blastopore lips. The ultimate result of this was the formation of **conjoint (Siamese) twins**; two newts with a fused body in the belly (the ventral). The blastopore lip is important in triggering the formation of the embryonic axis, and when it is additionally transplanted, it can drive a second embryonic axis that organizes a secondary embryo body plan. The blastopore lip was then named the '**organizer'**, due to its capacity to convince the cells surrounding it to make an embryo. Discovery of 'the organizer' was recognized with a Nobel Prize in 1935 to Hans Spemann.

Molecular mechanisms of the organizer

How does the organizer form? For more than fifty years after those impactful experiments, scientists tried to figure out what chemicals are secreted from the organizer to convince the cells in forming an embryo axis.

The molecular evidence came quite late. Establishment of the organizer is dependent on two signalling gradients, the first of which we already brought up, Wnt. The high Wnt signalling in the grey crescent, because of the cortical rotation, creates a side-to-side gradient that specifies the future dorsal side of the organism. From the vegetal axis (more or less perpendicular), there is another signalling pathway being turned on. The vegetal factors including the VegT transcripts turn on Nodal related proteins from the "bottom" (vegetal) to the "top" (animal), while the Wnt signalling in the grey crescent support expression of Nodal-related proteins as well. In sum, Wnt signalling and Nodal signalling centres at their intersection designate the dorsal vegetal cells named the "Niewkoop centre". Niewkoop centre induces neighbouring animal cells to express BMP and Wnt repressors, which makes those "the organizer". This also explains the earlier pursuit of the organizer, which looked into metabolites and activating signals, were in vain. It turns out that the organizer does not secrete signalling molecules, rather it secretes **signal inhibitors**.

Let us detail how those inhibitors turn on in a spatially constrained fashion. Genes such as *siamosis* and *twin* are in repressed state when the Wnt/β-Catenin signalling is inactive. Once the β-Catenin gets nuclear localized, it turns on those gene products which act as transcription factors. Alongside Smad2 downstream of Nodal signalling cascade, those transcription factors then turn on organizer genes: *e.g.*, *chordin*, *noggin*, *goosecoid*. Prior to organizer specification, embryo has BMP signal active throughout. On the side of the organizer *wnt8* establishes a gradient designating dorsal fates from the ventral fates. Organizer zone expresses many BMP/Nodal and Wnt inhibitors *e.g.*, *chordin*, *noggin*, *frizzled-b* (*frzb*), *dickopff* (*dkk*), *follistatin*, and *cerberus*. To summarize, right next to the Wnt and Nodal high intersection zone, a little piece of tissue, the organizer, fights off those two signalling centres as well as the global BMP signal across the embryo.

The formation of the D-V axis depends on the inhibition of two cellular pathways: 1- Wnt, which is at its highest next to the organizer and 2- BMP, which is diffuse throughout the embryo prior to inhibition. Once the organizer is specified, cells within this region are activated for the expression of inhibitors. The consequence of this is restriction of Wnt and BMP signals into spatial gradients: Wnt inhibited by Frzb and Dkk; BMP inhibited by Chordin and Noggin.

How is it that these inhibitors establish a BMP gradient within the embryo? We will look at the Chordin inhibitor protein, which is secreted by the organizer. As Chordin is expressed by the organizer located at the presumptive dorsal side, it gets secreted and interacts with the freely diffusing BMP ligand. Chordin binds to the BMP ligands, and this protein-protein complex has a very high diffusion coefficient, causing BMP to diffuse away from the organizer and towards the opposite end of the embryo. This shuttles the BMP ligands to the presumptive ventral side away from the organizer, creating a BMP gradient like a shuttle bus ring. Similar mechanisms are at play to switch the Wnt signalling gradient along the presumptive dorso-ventral axis.

The cartesian coordinates in perpendicular axes

How do we know the D-V axis is a real coordinate system, rather than a mere polarity event? In other words, does it have any **positional information**? To investigate whether the D-V axis encodes positional information dependent on the concentration of inhibitors, scientists titrated the BMP inhibitor Noggin in experiments. Researchers began with UV-irradiated embryos to halt the cortical rotation. Those UV-irradiated embryos only formed a belly piece: *i.e.*, a ventralized embryo missing any organizer activity. When those embryos were injected with *noggin* mRNA in titrated doses, in a dose dependent manner researchers could restore the embryonic development, *i.e.*, dorsalize the embryos. As Noggin decreases the BMP signal, the tissues were able to commit into more dorsal fates. However, when too much *noggin* transcripts were introduced to the embryo, the cell mass developed only into the anterior-dorsal structures. These experiments revealed that, in a dose dependent fashion, what the organizer secretes is able to determine both the anterior-posterior, as well as the dorsal-ventral structures. Therefore, the gradients formed by the organizer have positional information that instruct development of the embryo. Embryos have indeed Cartesian coordinates!

We had noted that the BMP signalling is everywhere: the default state for the early embryo prior to gastrulation is BMP signalling ON. What would happen if the global BMP signal persisted? The resulting mass becomes entirely ectoderm and epidermis. Conversely, when BMP is entirely depleted from the embryo, it commits into neural fate in entirety. Recall from earlier chapters that the induction of neural ectoderm was from the midline structure, the notochord. The notochord forms from the organizer. Proximity to the notochord, and therefore the organizer, results in lower BMP signal due to the secretion of inhibitors and

initiates neural differentiation. In the absence of BMP repression, the default of the ectoderm is to differentiate into epidermis (skin tissue).

While we have highlighted how the D-V axis is established so far. It must have become evident that the formation of the A-P axis is not separate or independent from the D-V axis establishment; the *noggin* RNA experiments detailed above made this clear. Wnt gradient before gastrulation was essential to specify the organizer and establish the D-V axis. What Wnt signalling does next, after gastrulation begins, is forming the A-P axis by localizing to the posterior side of the embryo. Wnt has two consecutive roles: first is defining the dorsal side and handing that D-V axis specification task over to the BMP gradient. Wnt gradient then specifies the A-P axis by localizing to the posterior end. Target genes of early Wnt signalling *e.g.*, *nodal-related*, *dkk-1*, *cerberus*, *noggin*, *chordin*, contribute to the dorsal mesoderm specification, organizer formation, and D-V patterning via the BMP gradient. Late Wnt signalling genes *e.g.*, *cdx, msx*, *meis*, *vent*, specifies the ventral and posterior mesoderm, contributing to the A-P patterning. The anterior-posterior axis is defined by late Wnt signalling, Nodal signalling, and later on Fgf signalling, which gets turned on by Nodal, is also recruited.

Back to the "Plan d'unité"

How does the molecular understanding change our perception of Saint-Hilaire's argument for a conserved body plan, and can it address why we see certain anatomical axes flipped? To answer this, we will look at the fruit fly development as an example for arthropods. Specific to arthropods, we know that the Dorsal protein (a transcription factor) is maternally deposited and has a gradient on the ventral side of the embryo due to symmetry breaking happening early in the oocytes. If Dorsal is absent, the whole embryo becomes dorsalized, and if it is no longer polarized but globally present, then the whole embryo develops ventral organ structures. In this way, Dorsal operates similarly to the organizer by playing a role in the formation of the dorsal-ventral axis. Unlike the development we have seen in amphibians and vertebrates in general, the BMP homolog in fruit flies, Dpp, is expressed at the dorsal end of the larvae. This creates a gradient where Dpp is higher in the dorsal side, and lower by the ventral side, opposite to how the BMP is oriented in vertebrate species. Dpp inhibitors such as the Short-gastrulation (Sog; Chordin homolog) protein localize BMP to the dorsal end. While the underlying concepts establishing the gradients are similar, and Dpp gets shuttled off to

the dorsal in a similar fashion by Sog, the localization of the BMP to the dorsal end flips the orientation of neurological development and the dorso-ventral body axis of arthropods in comparison to vertebrates.

"The organizer" across the vertebrates

We revolved around the amphibian development while explaining majority of findings discussed in this chapter. How generalizable this discussion is for the development of other vertebrate species and beyond? To this end, we will expand from the "organizer" of amphibians to the "organizer" of zebrafish. We know that in amphibians the organizer forms due to the high Wnt signal and high Nodal signal intersection at the Niewkoop's centre. The secondary body axis in amphibians is defined by *siamosis*, a gene targeted by the Wnt pathway, whose homolog in zebrafish is *dharma*. In zebrafish, we see that another cellular structure, referred to as the shield, acts as the organizer and defines the embryonic body axis. Similar transplantation experiments conducted in zebrafish, moving the organizer of one embryo into another, likewise results in conjoint twin phenotype driven by two organizers.

What does the "organizer" concept look like in amniotes, animals that develop within additional amniotic sac, *e.g.*, birds and mammals? While the geometry of early embryo varies drastically in between amniotes, they all possess an organizer tissue that is referred to as "node". Using chicken development as an example, let us recall the primitive streak that forms during gastrulation. Through the primitive streak, the cells delaminate, rather than invaginate, to form the embryonic germ layers. As the collective cell migrations of gastrulation continue, the base of the streak regresses towards where the tail will form, and the "start" of the streak forms the anterior side. The regressing end of the primitive streak is referred to as the node, and the cells left behind after delamination form the notochord.

Let us compare the gastrulation of a mouse to a chicken. The case for mouse embryonic development is peculiar, where the embryo is shaped like a cup, possessing an internal cavity where the endoderm is initially outside, and the ectoderm is inside the cup. The endoderm shifts into the innermost layer later during gastrulation, where the whole embryo undergoes a complete inversion and its contents flip. While this is quite different than what is observed for other amniotes, the mice as well develops the primitive streak along its future A-P axis. The primitive streak forms and the regression of the streak, similar to the case for birds,

specifies the site wherein the organizer, node, will form and specialize into the notochord. While the developmental geometry that leads to the formation of notochord differs in between the examples so far discussed, a similar kind of organizer forms during gastrulation.

For the **Spemann organizer** in frog embryos, we saw that the ectoderm closer to the organizer forms the neural tissue and ectoderm that lies further away forms the epidermis. The story is similar for fish, with the difference being that the fish embryos gastrulate over a large yolk and gastrulation closes the blastopore lip. However, again, the ectoderm that lies closer to the **Shield organizer** differentiates into neural tissue. For chicks, we see the same scenario, except the organizer, **Hensen's Node**, now lies at the regression of the primitive streak. As the streak continues to regress in the chick embryo, it will give rise to the flanking mesoderm and, unique to amniotes, extra-embryonic tissue. This extra-embryonic tissue is similarly formed by **the node** in mice. Overall, we observe a convergence in developmental processes at the stage of organizer formation, highlighting developmental constraints on evolutionary flexibility.

Regardless of the name and organism, the organizer has the unique function of expressing both Wnt and BMP inhibitors and as a consequence will dorsalize the embryo. When present, the organizer induces the formation of the embryonic axis. In 1934 Conrad Waddington, a founding figure of developmental genetics, assessed how the organizer of one species could affect the development of another species. Grafting the organizer tissue from chick to rabbit embryos (**Figure 9.2**) Waddington noted: "Owing to its comparatively slight tendency to stick to the operating knives, chick tissue can fairly easily be grafted into the embryonic shield of the rabbit. In two cases, grafts of the chick primitive streak have succeeded in inducing the formation of neural tube from the overlying rabbit ectoderm." The organizer surprisingly, was able to function **interspecies**, inducing the neural fate across species. This was recapitulated in another set of experiments, where chunks of tissues from various positions of the chick embryo were explanted and sandwiched between the ectoderm of a frog embryo. The frog ectoderm began expressing neural fate markers, however, only when the graft was coming from the chick's Hensen's node (**Figure 9.2**).

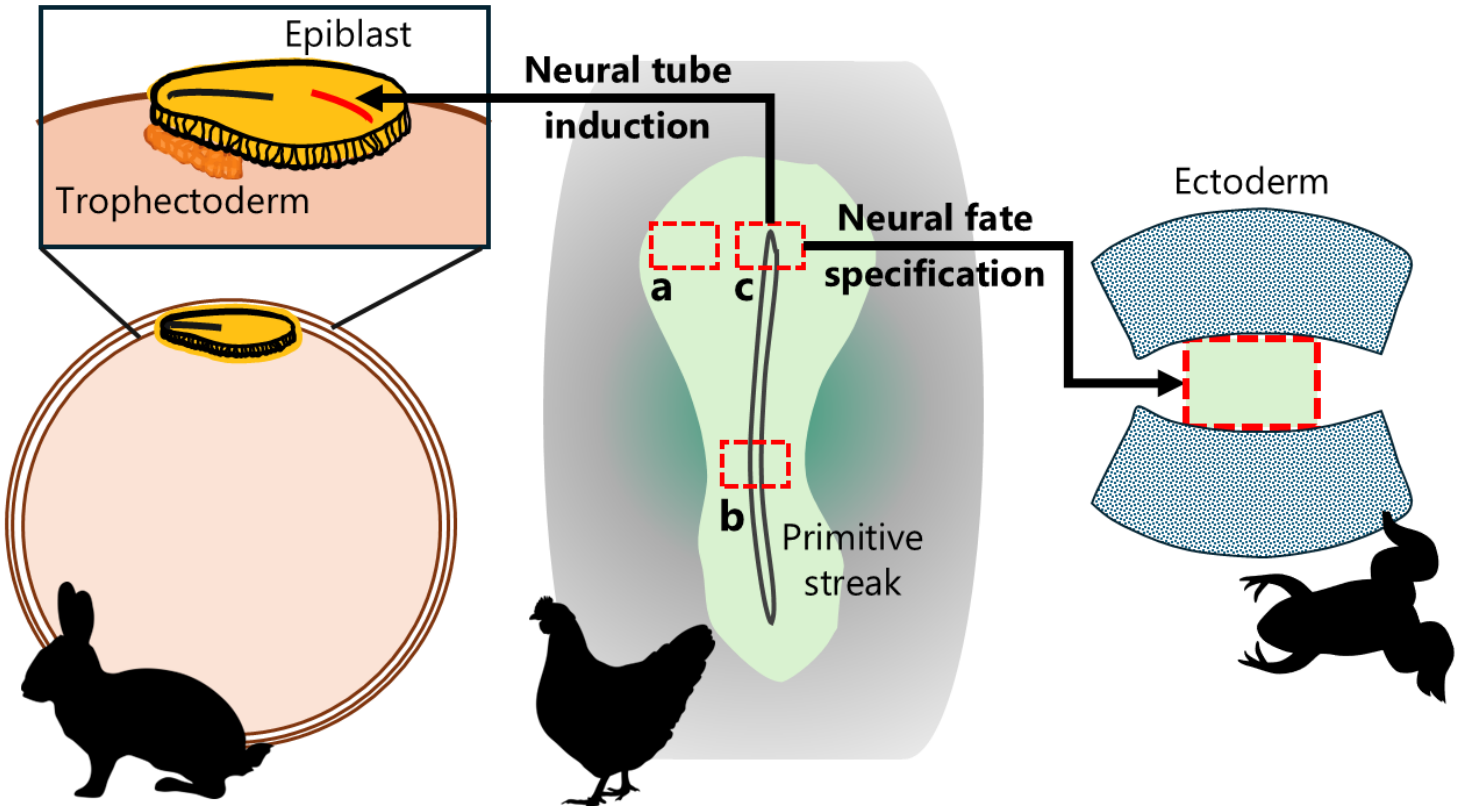

Figure 9.2 | The chick organizer (**middle**, c, but not other tissue explants, *e.g.*, a or b) can induce neural tube once grafted in rabbit embryos (**left**). Likewise, if sandwiched in between the frog ectoderm tissue, the chick organizer induces neural fate in the Xenopus surface ectoderm (**right**).

Embryos come into a 3-D formed structure with anterio-posterior, dorso-ventral, and medio-lateral axes as the gastrulation proceeds. The organizers of each species, expressing Wnt and BMP inhibitors, are essential for that specification. In vertebrates, the organizer dorsalizes the embryos and establishes the A-P axis followingly, and itself eventually differentiates into the notochord. Specifying the embryonic midline

establishes the medio-lateral axis as well. Altogether, organizer puts the embryos into a similar body plan and structure: "the phylotypic stage" of neurulation, which will be our next station.

FURTHER READING

Geoffroy Saint-Hilaire: a visionary naturalist. H. Le Guyader, **University of Chicago Press**, Chicago, 2004. https://archive.org/details/geoffroysainthil0000legu.

The organizer: What it meant, and still means, to developmental biology. J. Slack, 2024, **Current Topics in Developmental Biology.** DOI: 10.1016/bs.ctdb.2023.12.001.

Chapter 10: THE PHYLOTYPIC STAGE OF VERTEBRATE EMBRYOS

"I HAVE TWO SMALL EMBRYOS PRESERVED IN ALCOHOL THAT I FORGOT TO LABEL. AT PRESENT I AM UNABLE TO DETERMINE THE GENUS TO WHICH THEY BELONG. THEY MAY BE LIZARDS, SMALL BIRDS, OR EVEN MAMMALS."

– Karl von Baer, 1828.

In Chapter 2, we had introduced a special stage of early vertebrate embryogenesis, the phylotypic stage, during which all vertebrate phylum embryos look alike. This is a similarity going beyond mere morphology, setting the development program under similar constraints. What minimizes the divergence between early embryonic geometry and sets the phylotypic stage? The organizer is key here with actions that can be categorized into four. First, organizer secretes BMP and Wnt inhibitors, dorsalizing the embryos alike. Next, the organizer tissue itself begins differentiating into the notochord marking the embryonic midline. Third, organizer/notochord induces the neural ectoderm atop from the surface ectoderm by inhibiting BMP signalling. Lastly, this embryonic midline becomes surrounded by the paraxial mesoderm laterally which gradually give rise to tissue blocks known as **somites**. Those two events, neurulation (fate commitment and patterning of the nervous system) and segmentation (antero-posterior axis into somites) of embryo, are the landmarks of the vertebrate phylotypic stage.

Segmentation

The process of vertebrate embryo segmentation, somitogenesis, begins at the posterior paraxial mesoderm. As the node is regressing towards the tail and axis elongation is taking place, pairs of tissue blocks are forming bilaterally (flanking to the left and right sides of the A-P midline). Those bilateral pairs of paraxial mesoderm blocks which **sequentially segment** from the presomitic mesoderm (PSM) in the tailbud are called somites. Forming pairs are numbered based on their order of segmentation; S1 is the first pair in the very

anterior, S2 the second, so on and so forth. The rhythmicity of this sequential segmentation is coordinated by a molecular clock.

The role of somites, as repetitive blocks, is critical for a developing vertebrate embryo. Somites provide positional information for organogenesis and pattern the embryonic axis. Each somite pair differs from the other, and as they are not identical, they can induce the cells that neighbour them to commit to different fates. For example, in a zebrafish embryo, cells that are adjacent to the S4 somite pair differentiate into the pectoral fin (homologous to forelimb of a tetrapod), while the cells adjacent to the S17 pair end up becoming the hind limb homolog. Differentiation of cells into organs, limbs, the peripheral neural system, and circulatory system is guided by segmented somites. That way, somitogenesis, hence the phylotypic stage, acts as a bridge between gastrulation and organogenesis.

Notably, two key structures that guide differentiation of neighbouring tissues, the **notochord,** and the **somites**, are both embryonic tissues not found in the adult organism. The notochord vacuolizes and mostly disappears, leaving behind only a remnant piece in the centra of the vertebrae. But somites are absent due to a different reason; they eventually further differentiate. Somite cells contribute to various adult tissues including vertebral column, back muscles and epidermis, connective tissue in vertebrates as well as brown adipocytes, cells important for thermal regulation, in mammals. They also contribute cells to the outgrowing limb buds.

Mechanisms of somitogenesis

The periodicity of segmentation and total number of somite pairs forming are both **versatile**, that they vary between species, and **robust**, are tightly species-specific. For instance, zebrafish is one of the fastest examples among model organisms, with pairs being segmented every 25 minutes, while humans take around 5 hours to form a pair of somites. The segmentation period is 90 minutes for chick embryos, 120 minutes in mice, and ~100 minutes in corn snake. Total number of somite pairs forming is 32 in zebrafish, 52 in chick, 60 in mice, and more than 310 in corn snakes. The figure for human embryos is a bit ambiguous that it segments ~45 somite pairs, however, end up having ~40 pairs due to tail truncation and loss of few somite pairs later in development.

As the previous section laid out, the position of a somite pair, *i.e.*, how big somites are, is very important in the developmental patterning of an organism. This depends on where in the presomitic mesoderm, the cells "decide" to become somites. This was investigated with a series of classical embryology experiments in chick embryos. An anterior (closer to segmented somites than the tail bud) portion of the presomitic mesoderm of the chick embryo was excised and grafted back to the embryo in a A-P flipped orientation. Researchers observed that the order of segmentation of this "flipped" block of tissue was also inversed. If we align the chick embryo with head at the top and tail at the bottom, the somites once at the "top" of the PSM were developing first from the bottom of the inverted block along the A-P axis. Additionally, it was seen that **only** the excised region observed this flipped order of segmentation. This ultimately revealed that the order of segmentation is defined early in the PSM; the cells of anterior PSM that have yet to become somites **are already committed** to segment at specific positions and times. This temporal lapse from commitment to the segmentation is due to the epithelialization process of mesenchymal cells surrounding future somite boundaries. The molecular program takes time:

1- genes need to be turned on-off for polarity and epithelial commitment,

2- the cells need to express ligands and receptors to push each other and form the intersomitic furrows,

3- cytoskeleton needs to be remodeled, and

4- proteins need to be secreted into the extracellular space for basal membranes of the epithelium to attach.

We previously mentioned that a molecular clock determines the timing of somitogenesis. How so? The transcription factors we brought up in Chapter 3 while covering the Notch signalling, hairy-enhancer-split related (Her/Hes) proteins, are conserved molecular oscillators found in each vertebrate species studied so far. Intriguingly, we observe that every cell in the PSM experiences Her/Hes oscillations at **every** segmentation cycle; however, only a single batch commits to forming a somite pair. What localizes the cellular decision-making process? An additional layer of control through Fgf/Wnt signalling gradients sourced from the tail bud leave that only certain cells respond to the "clock" of each cycle.

Recent studies in zebrafish clarified that the molecular clock has two functions carried out consecutively: segmental commitment and polarization of every somite pairs. In the first process, segmental commitment,

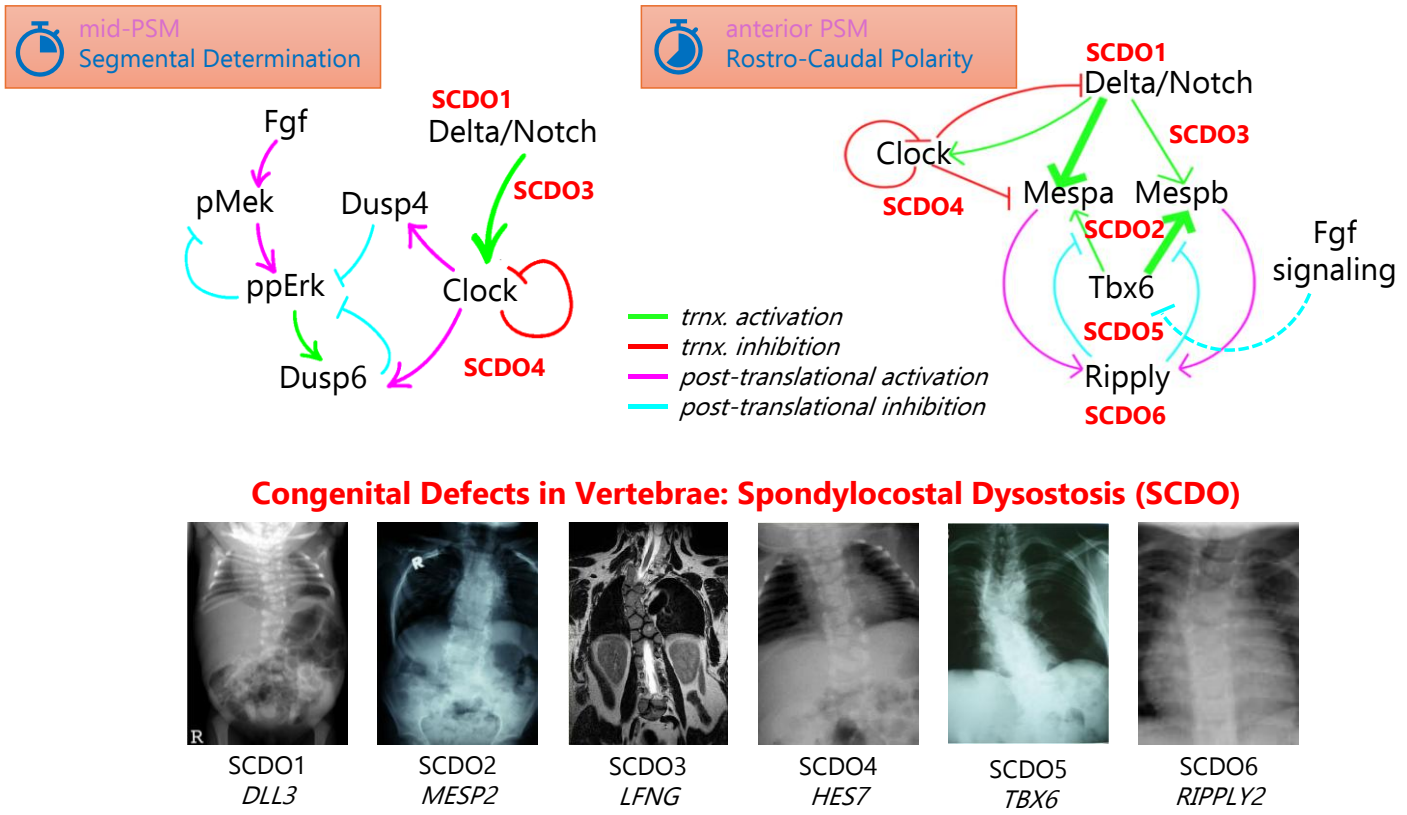

Figure 10.1 | Molecular (transcriptional and post-translational) regulatory network functioning for two consecutive roles of segmentation clock in patterning somites as revealed in model organism studies and six genetically associated congenital defects of vertebrae in patients (Spondylocostal dysostosis). Radiographs are adapted from: Sparrow *et al.*, Human Molecular Genetics, 2013; McInerney-Leo *et al.*, Human Molecular Genetics, 2015; Lefebvre *et al.*, Clinical Genetics, 2016; and Turnperry *et al.*, Gene Reviews, 2017.

the molecular clock interacts with the Fgf signalling gradient to control the sequential segmentation rate and place. Once certain cells commit to making somites, each somite will further become polarized into the rostral ("head facing") and caudal ("tail facing") compartments. The anterior side of the somite will be different from the posterior side of the somite. This also is carried out by the segmentation clock acting in the anterior PSM as a thinning stripy expression and a gene regulatory network linked to the clock. Although these regulatory networks are mapped through model organism studies, mutations to the genes acting result in congenital vertebrae defects in human patients (Spondylocostal dysostosis, SCDO). Disease-linked proteins identified so far include DLL3 (Delta-like three; Notch, SCDO1), MESP2 (target of Fgf and Notch, SCDO2); LFNG (Lunatic fringe; Notch, SCDO3); HES7 (Her/HES transcription factor, SCDO4); TBX6 (target of Fgf signalling, SCDO5), and RIPPLY2 (target of MESP, SCDO6) (**Figure 10.1**).

How do we know that each somite segmented specifies a certain position along the embryo axis? Some classical experiments in chick embryos cast light on the process. Embryologists dissected an anterior PSM tissue out of an older stage donor embryo, engrafted it into an earlier stage host embryo's anterior PSM and allowed the host embryo to develop. Developing chick embryos formed thoracic vertebrae (vertebrae with ribs) in the place of their cervical (neck) vertebrae indicating the positional identity coded and preserved along the A-P embryo axis.

Neurulation

In the morphogenetic principles and gastrulation chapters, we have discussed that the neurulation begins by the formation of a neural sheet from ectoderm – induced by the notochord along the mid-embryonic axis – which then invaginates and fold into a tubular structure: the neural tube. In upcoming chapters, we will dive deep into how the neural tube is fashioned into specific neuronal fates. For now, let us expand on the neurulation process from the perspective of the coupling between underlying mesoderm (paraxial, segmented somites) and the neural ectoderm. Ice-breaker for the discussion is a peculiar cell type: the **neural crest cells.** Neural crest is an evolutionary innovation in vertebrates. It is useful to remember that ectoderm's default state is neural fate commitment, but high BMP signal across early embryo prevents them to commit.

Organizer/notochord fights off the ventralizing BMP source to clear a space atop (dorsal to notochord) for neural induction of the ectoderm sheet along the midline. Cells laterally flanking the neural sheet are "in-between" the high and low BMP signal (**Figure 10.2**). This just enough BMP signalling deems neural crest progenitors unable to form either the neural tissue or the epidermis. Once the neural tube folds, neural crest cells sitting between the neural tissue and the outer skin become **highly migratory** delineating atop the neural tube. These cells travel during development and play a multitude of roles. Some commit to neural fates, while others contribute to the gut or turn into muscle, cartilage, and bone. The fact that human faces are different from one another, and we can recognize each other

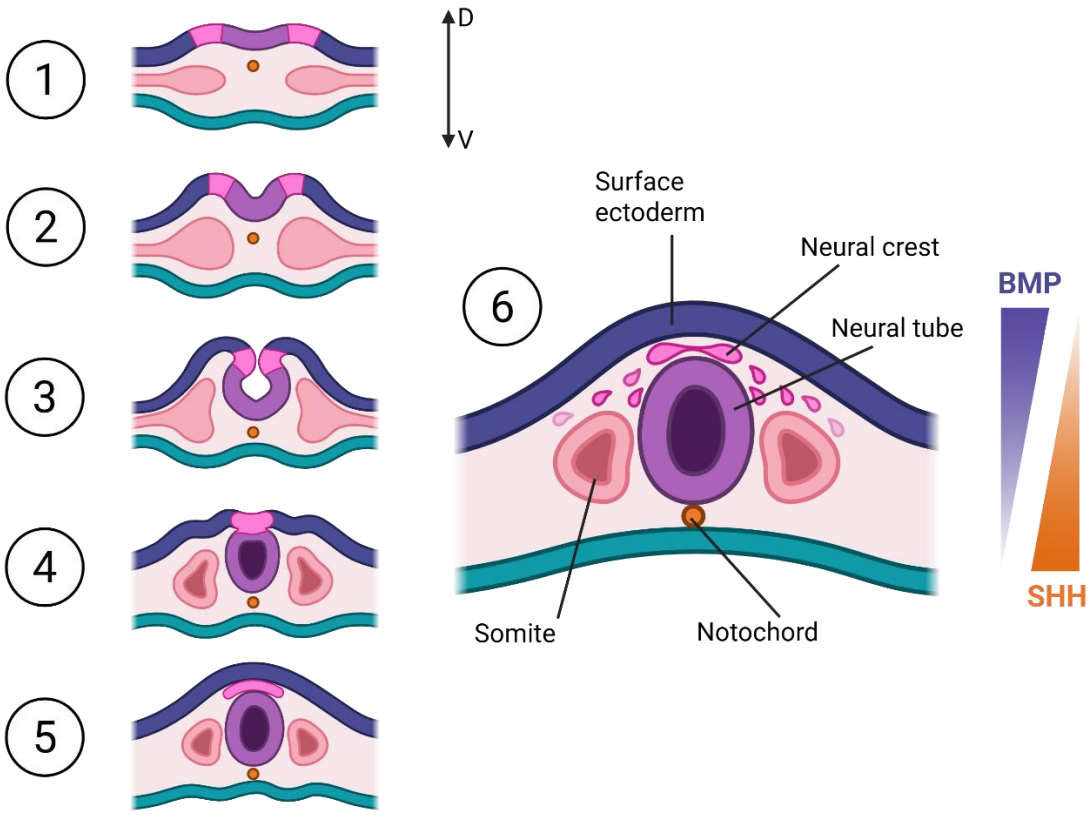

Figure 10.2 | Specification of neural crest cells (pink) flanking the neural sheet (purple) by receiving medium levels of BMP signal from the surface ectoderm (dark blue). Neural crest cells delaminate and migrate after neural tube formation with guidance from somites.

easily boils down to those cells. Their partially stochastic (unpredictable) migration for every developing human embryo, and later differentiation into craniofacial muscles, or ossification into cranial bones, gives rise to this recognizable diversity among us!

Timing and directionality of neural crest migration is influenced by the maturity and polarity of their adjacent somite. The gradient of cellular maturity is first established along the anterior to posterior axis of an embryo. Somites closer to the anterior pole of the A-P axis segment first, and this results in anterior somites being more mature than those in the posterior. As the somites mature, the neural crest cells first track over rostral halves of somites during their lateral migration from the midline. Later in development, the axons from the neural tube also extend over the rostral somite halves. Eventual patterning and connection of spinal cord with the peripheral nervous system arise as such. A transplantation assay proves useful here to understand the process. If one removes an individual somite pair and returns it to the organism in flipped orientation, the growth of the axons in that region also follow the flipped track. From this, we see that the site of axon outgrowth is not a feature of the neural tissue, rather the information comes from the somites beneath it. However, it is not that somites are required for axonal growth. If all somites are removed, there is no patterning or restriction, and axon outgrowth happens throughout the whole organism.

FURTHER READING

Cellular and molecular control of vertebrate somitogenesis. Y. Miao, O. Pourquié, **Nature Reviews Molecular and Cell Biology**, 2024. DOI: 10.1038/s41580-024-00709-z.

Periodic inhibition of Erk activity drives sequential somite segmentation. M.F. Simsek, *et al.*, **Nature**, 2023. DOI: 10.1038/s41586-022-05527-x.

Gene-regulatory interactions in neural crest evolution and development. D. Meulemans, M. Bronner-Fraser, **Developmental Cell**, 2004. DOI: 10.1016/j.devcel.2004.08.007.

Chapter 11: DEVELOPMENTAL GENETICS

"Все счастливые семьи похожи друг на друга, каждая несчастливая семья несчастлива по-своему.

All happy families resemble one another; each unhappy family is unhappy in its own way."

– Leo Tolstoy, in Anna Karenina, 1877.

In the very first steps of this voyage, we had asked the question of "how an embryo is built" and elaborated that the cells are the master builders of the embryo. In this vein, we covered what cells were doing, how they specify into certain fate commitments and how changes in their behaviour allow for embryonic development. We learnt about the changes to cell motility and migration, the importance of this for the collective tissue movements during morphogenesis. In this process, cells have to communicate with each other, and we have covered the bulk of the cell signalling pathways activated during development. Many landmark events in development are realizations of cellular decision-making processes. That ongoing discussion brings us to a particular perspective of development oriented around the genes now. The role of genetics during development can broadly be referred to as **developmental genetics**.

From our daily lives, we know that genes are important and central to that question we have been digging answers for so far. Consider the event of monozygotic (identical) twins, which happens when a fertilized egg is split during early cleavage events into two. This splitting can happen from time to time, and these individual pieces develop into genetically identical but separate embryos. We can look back to Driesch's experiments, where he split the 4-cell stage embryo into individual cells and saw four sea urchins forming. The concept of monozygotic twins is quite similar. Recall, the fertilized egg is a genetic mixture of both paternal and maternal DNA, and as the cleavage splits the fertilized egg into two masses, the cell division and duplications lead to two siblings. And we see that many features that form between these two siblings are almost identical, down to the point of how the neural crest cells migrate and shape the faces of these individuals. However, we must keep in mind that, for this case, we are discussing *in utero* development, where not only is the genetic content the same between these individuals but also their environment. And with the environmental effects, as twins

age, they might begin to diverge in looks. Long story short, the very presence of identical twins shows us that genes are important and central.

Birth of modern genetics

To talk about the importance of genes for development, we must first go back in time to the very foundations of modern genetics. Attending animal breeders' convention, Thomas Hunt Morgan was upset with the state of Mendelism, which was what genetics used to be referred to as. In the 1900s, hereditary factors were used to explain away observations in development, why certain breeds turned out a certain way, but no one knew what these factors were. Consequently, many of these explanations used a combination of facts and "alternative" facts, turning the field into a pseudoscience. Morgan wrote in 1909: "In the modern interpretation of Mendelism, facts are being transformed into factors at a rapid rate. If one factor will not explain the facts, then two are invoked; if two prove insufficient, three will sometimes work out. The superior jugglery sometimes necessary to account for results may blind us, if taken too naively, to the commonplace that results are so excellently 'explained' because the explanation was invented to explain them…"

To systemically tackle the issue, Morgan conducted work on fruit flies by breeding thousands of flies and investigating the phenotypes of their progenies. Those efforts led to identification of the first mutational phenotype for animals, which was a white eye phenotype (*white*) restricted to male flies. Given that this mutation appeared linked to the sex chromosome of the flies (XO), he assessed the inheritance of this mutation through experimental crossings. He crossed a wild-type (red eyes) female with a *white+* male, generating an F1 heterozygous population. If the *white* phenotype were recessive and inheritable, he expected to find that, in the second crossing of these heterozygous offspring, the white-eyed phenotype would reappear. This is exactly what he saw; he observed a ration of 1:3 of white-eyed animals in the F2 generation. With these observations, T. H. Morgan's work supported Mendel's observations of inheritable traits from the plants and brought vague hereditary factors of Mendelism into "physical" gene loci.

Together with his group in Columbia University, in their famed 'fly room', Morgan was able to further screen for different types of sex-restricted mutations and find others that impact eye colour, body and wing development. Notable mutations his lab characterized include yellow (where female flies have a yellow

colouration), eosin (which affects eye shape), miniature, black, rudimentary, to name a few. All these mutations segregated with the sex chromosome, some being specific only to females and others only to males.

This data collected over the years characterizing various sorts of sex-chromosome linked mutations in flies was used by Alfred Sturtevant, an undergraduate student in Morgan Lab. Sturtevant postulated that, if the heritable factors were organized within a chromosome, in what we'd refer to today as a locus, we should see proximity-dependent frequencies of crossing-over events. For example, if gene A and gene B were further away from each other, we would expect them to segregate away from each other more frequently than two genes closer together in their loci. Using the logbooks of the 'fly room', he tried to mathematically sort the order of the mutations. And he was successful. The results of his work were published in a single-author paper, and "*... form[ed] a new argument in favor of the chromosome view of inheritance, since they strongly indicate[d] that the factors investigated are arranged in a linear series, at least mathematically*." This landmark discovery which devised the first gene mapping calculation was the birth of modern genetics. However, at this stage it was still difficult to do embryological experiments with mutants, DNA was not discovered, central dogma of biology was unknown. It took some forty fifty years for those seedbed experiments to flourish into a whole field of inquiry: field of developmental genetics.

From genotype to phenotype?

In the fly room, scientists learnt that a variety of mutations can cause unique phenotypes; there are white-eyed flies, miniature flies, black flies, all driven either by a single gene or a cooperative impact of multiple mutations in a locus. But how do we then understand what genotype gives rise to what phenotype? That $100 billion question of biology is not one we have a clear answer to. And this might be because it is the wrong question to ask. We shall keep in mind that a phenotype is the end result of certain or many cellular decision-making processes throughout the development. However, when we speak of genotypes, in the process of cell duplication events from the fertilized egg, the zygote, every single cell in the animal which will carry out all sorts of differing tasks and commitments has basically identical content. Our ability to map certain gene mutations to certain phenotypes, as we have seen in fruit flies, does not guarantee that the gene itself is causing something to happen. It rather merely proves that if that gene has a particular mutation, some cells in

the developing embryo will be unable to carry out that specific task. And that can deceive us that genotype makes phenotype. We can elaborate on this with a similitude: If a horticulturist does not water her garden for a week or so, plants in the garden will likely begin to die off and the patterns drawn with colourful flowers will begin disappearing. However, looking at this, if an outside observer concludes that the water is driving the pattern formation in the garden, that will be a dire inference! Likewise, genetic mutation studies may come prone to such hasty conclusions for embryonic development.

To clarify, let us go through two examples where single gene mutations may be linked with very observable phenotypic effects. Looking at Dachshunds, we can relate their most notable feature to a single gene. Their short legs appear to be caused by a duplication of the Fgf4 allele, as a result of selective breeding. With this extra copy, chondrocyte development prematurely stops the growth by accelerating the differentiation of the chondroblasts. Overall, fast differentiation results in short stature without much bone growth in this dog breed. Another example of how single gene mutation can drastically change phenotype, again from Dachshunds, is related to the long-haired variety of the breed. A loss of function, LOF, mutation in another Fgf ligand in that case, Fgf5, which is implied to control the regression stage of the hair growth cycle, is linked to the long-haired phenotype. Due to loss of function, the hair cycle basically keeps happening without the regression stage. At that point we know that the changes to a single gene's sequence or copy number may cause drastic phenotypic outcome. Sometimes! However, even in such seldom cases where single alleles are linked, it is hard to predict what genotype would result in what kind of phenotypic outcome. In many cases, this is the wrong question to begin with. When we consider those two examples in the light of what we have learnt so far about the Fgf signalling pathway, it is immediately clear that "why of the phenotype" is a challenging question. It is hard to predict how a cell would respond to an extra gene copy or LOF mutation in two genes coding for Fgf ligands belonging to very same FGF subfamily. Those genes are transcribed and translated into proteins which are secreted as a signalling ligand, *a priori*, binding to and activate same set of FGF receptors. How would the signal be adapted by the built-in feedback mechanisms or respond to the changes in incoming ligand levels? How would the development respond to heterozygous mutations of the same allele? Would a gene duplication or LOF mutation in another Fgf ligand or receptor allele result in similar phenotypes? The effect of gene mutations may sound clear with a monogenic phenotype, but the *why* is often less obvious.

A single gene may play a large phenotypic role, like Fgf in Dachshund characteristics discussed. However, most times, the case is not as straightforward and "what gene is regulating this process" might not be the correct question to ask. Why so? There is no blueprint nor a cookbook in the genome which gets read by the cells to build the structure, to prepare the recipe. Along the path of development, as a fertilized egg divides and multiplies, each cell makes decisions rather than "going from genotype to phenotype". These cellular decisions are made at a systems level, not at the genome level; and by the systems level we mean the information universe encompassing the cytoskeletal features, extracellular materials, cellular communications, and the physical forces between different cells.

The epigenetic landscape

Bringing up Conrad Waddington, a founding figure in developmental genetics, into discussion from 1950s will be beneficial. Waddington coined the –now famed and stretched– term "epigenetics" in his analogy of balls rolling down a hilltop to emphasize the factors beyond genotype. Imagine a cell as a ball rolling down a hill, which have valleys and peaks and as many balls alongside roll down their paths might bifurcate. In each cell, there is the same core genomic material, but epigenetic factors shape the paths a cell would go down. Eventual to the chosen paths, the cells become very different from one another. This "epigenetic landscape" puts the genotype to phenotype question into a cartoonish but accessible scheme. Alternatively, we could also imagine cells like mountaineers. As the cells are trying to make decisions along the way of their climb to build the embryo, genes are the material the cells have at their disposal, *i.e.*, the mountaineering gear for their trek. If those genes are not present, due to a deletion or mutation, we can see that the cells cannot accomplish their tasks. In the end case, we see the cells completing the task differently or abolishing a certain task altogether. So, when we say, "genotype causes phenotype", it is more of saying that if the cells cannot find certain tools at their disposal, they will not be able to carry out a certain developmental task during a specific period as it happens in the wild type. Depends on the nature of outcome, scientist would call this a loss of function or gain of function mutation.

Developmental genetic screenings

The understanding gained from the effects of selective breeding and Morgan's experiments in the early 1900s later coagulated with discovery of the "gene" code in the 1960s. The **central dogma** of biology was posited: genes are transcribed into RNAs which are translated into proteins, functional machinery of cells. Information flow was unidirectional (mostly; many exceptions were discovered later as happens for everything else in the science of life!). At that point, scientists thought they can carry-out extensive genetic screens and identify the genes important for early embryonic development.

The fruit fly gets segmented along its major body axis, anterior-to-posterior, into fourteen defined segments (3 cephalic, 3 thoracic, and 8 abdominal) which give rise to distinct parts of the adult fly. Christiane Nüsslein-Volhard and Eric Wieschaus thought this is an ideal organism for the goal. If they screened those segments, *i.e.*, what happens to them –how are they shaped, whether some merge or disappear– in that very early larval stage, they can identify the genes which are impacting the embryonic development. Those screens proved extremely successful, identifying many regulatory genes including the one discussed in Chapter 3: the hedgehog. Nüsslein-Volhard and Wieschaus shared the 1995 Nobel Prize in Physiology and Medicine for uncovering the developmental mechanisms of the embryo through genes, alongside a third scientist we will introduce in the next chapter, Ed Lewis.

While a screen might sound simple at first – you mutate the flies and observe the larvae – thorough genetic screens require a lot of clever planning. Their screen involved three generational crossings and specific male and female fly genotypes. Within their screen, they would select a tractable mutation in their chromosome of interest in males, for example, the *white* allele in the X chromosome. They would then use females carrying two mutations in the same chromosome: **Dominant temperature sensitive** (Dts) which are nonviable at higher (>29 °C) temperatures and **b** which is homozygous lethal without any observable pattern abnormalities in the larvae. The tractable males then were treated with the chemical mutagen EMS, causing a mutation in an allele (a*) and crossed with the (Dts/b) females. Next, the (a*/b) males are selected and crossed again with the (Dts/b) females. The second-generation larvae, grown in > 29°C, are only viable if they are either male or female heterozygous (a*/b). In their final crossing, male (a*/b) are crossed with female (a*/b) resulting

in either: (a*/a*), (a*/b), or (b/b) progeny. If a* was a developmentally important allele, the resulting larvae would not be viable but would exhibit a patterning defect.

When looking at the dead (a*/a*) larva, there are several features to look at: the morphology of the anterior or posterior ends, the segmentations of the body axis, the cephalic, thoracic, and abdominal segments can also be distinguished by cuticle formations and denticles.

We will go through two examples. The first group of phenotypes we will discuss are what Nüsslein-Volhard and Wieschaus categorized as **gap genes**; homozygous recessive, flies with a Gap gene phenotype have shifts or gaps through their segmentation. For example, flies can lose two of their three (T2 and T3) thoracic segments. In this case, the first thoracic segment (T1) is immediately followed by the first abdominal segment (A1). These observed shifts, or gaps, to the segments are the namesake for the phenotype, indicating cells need the gene of interest to carry out the localized patterning commitment for those absent segments. The second example is what arises from mutations in the **pair-rule genes**. When these genes are mutated, the resulting phenotype is skipping every other segment identity in an orderly fashion: as if fusing their certain segment pairs into single ones depending on the gene of interest. Eventually, mutations in pair-rule genes lead to a characteristic phenotype where the embryo develops with half the normal number of segments.

Developmental gene regulatory network for fruit fly segmentation

1. Maternal genes

To understand the role and mechanism of these patterning genes, we have to start from the beginning, the fertilized egg, and return to the maternally deposited factors. The story begins with the *bicoid* mRNA which is maternally deposited, held locally in the anterior end of the egg, and gets translated to Bcd transcription factor after fertilization. The polarized RNA gradient results in a Bcd gradient throughout the egg. In the absence of this maternal factor, we see that the larva loses anterior structures, having no head, no T1-T3 segments, and no A1-A5 segments either. When researchers performed experiments by transferring the cytoplasm of a wild-type egg into a *bicoid* mutant egg at different positions, the wild-type cytoplasm transferred to the presumptive anterior end of the egg was able to recover some of the anterior structures.

However, if the cytoplasm from anterior Bcd high region is transferred to the centre of the *bicoid* deficient egg, a headlike structure formed in the middle of the larvae surrounded by a mirrored image of patterning: to the left and right of this head are thoracic segment in order (T1, T2, …). These results confirm that maternal Bcd plays a necessary and sufficient role in anterior development.

The Bicoid transcription factor mainly plays two roles: it activates zygotic *hunchback* (Hb) expression and inhibits the translation of the *caudal* maternally deposited mRNA which is evenly distributed throughout the egg. Notably, although Bcd activates Hb expression, *hunchback* mRNA is also present as a maternal factor evenly diffuse in the egg. Prior to the maternal-zygotic transition (MZT), the result of Bcd expression at the anterior end is higher Hb and lower Caudal protein expression, relative to the posterior. Hunchback acting in the posterior is detrimental for development. A fourth maternal gene: *nanos* at the posterior end of egg (opposite to *bcd*) comes here to the rescue! Nanos is not a transcription factor. Its posterior to anterior protein gradient rather prevents Hb translation in the posterior. Mutants for both maternal *hunchback* and *nanos* develop normally! The oocyte came with four maternal RNAs: *bicoid* localized in the anterior and *nanos* localized in the posterior end, *hunchback* and *caudal* deposited homogenously along the axis. End picture at early cleavage stage is four protein gradients: Bcd and Hb from anterior to posterior; Cad and Nos from posterior to anterior.

2. Gap genes

Following MZT, three maternal transcription factors turn on the zygotic *gap* genes: *hunchback*, *giant*, *krüppel, knirps,* and *tailless*. The last one also gets input from a terminally restricted RTK signalling, localizing it to the very ends of the embryo. Those five genes are coding for transcription factors (TFs) with localized expressions along the embryo and mutations to any result in the "gap phenotype" we previously introduced. In this regulatory network, Bcd activates the expressions of *knirps* (*Kni*), *hunchback* (*Hb*), *giant* (*Gt*), as well as the *tailless* (*Tll*) gene in the anterior, while Cad activates the expression of only *Gt* in the posterior. Activation of *krüppel* (*Kr*) as a band crossing mid-embryo is peculiar. Hb is the activator for *Kr*, however, when over-expressed same TF turns into a repressor of *Kr* expression. This is thought to be because Hb dimerizes more over a certain level and the dimers block polymerase activity instead of working as an activator. Since Hb is also turned on zygotically in the anterior by Bcd, this anterior Hb overexpression prevents Kr to expand

anteriorly, limiting it into a sweet spot in the middle of the embryo. If you mutate Bcd so Hb levels go down within this sweet spot, Kr can invade all the way into the anterior ends. Further cross-regulation between the gap gene network TF pairs establish a key network motif: **toggle switches**, where the TF pairs repress each other's expression in a tug-of-war like fashion. Toggle switches ensure expression of each gap gene along the embryo axis in their own territory.

a. Giant (Gt) and Krüppel (Kr) toggle switch

Gt is turned on by Bcd and Cad, both in the anterior and the posterior, but excluded from the posterior and anterior ends of the embryo with repression by Tll; whereas Kr is turned on by Hunchback in the middle of A-P axis. The Gt – Kr toggle switch defines exclusive zones for Kr and Gt expression.

b. Knirps (Kni) and Hunchback (Hb) toggle switch

Knirps, activated by Bcd, is repressed by Hb and represses it back. As both the maternal and the zygotic Hb are enriched at the anterior half of the embryo, there are lower levels of Kni in this region. Kni expression zone sandwiches the Hb from both sides.

3. Pair-rule regulatory network

Pair rule genes are downstream targets of the gap gene network, defining every other segment of the fruit fly as mentioned earlier. For instance, *even-skipped* gene is only expressed in the odd-numbered segments; conversely, *fushi tarazu* is only in the even-numbered segments. Mutations of those genes result in loss of those segment identities; with half of the segments lost, pair-rule gene mutants will have 7 larger, rather than 14 regular, segments.

Along the segmentation of the fruit fly body axis, the pair-rule genes act as the next control panel under the gap genes to establish finer patterning. How is that succeeded? We will go through one example, for *eve* stripe 2, which is expressed solely in the third segment under a tight regulation. If we look at the promoter region upstream of this gene, we can see binding sites for Bcd, Hb, Gt, and Kr. Bcd and Hb are activators of the gene, while Gt and Kr are repressors. Looking back at the gradients of all these transcription factors within

the larvae, it is only in the third segment where we see sufficiently high Hb and Bcd levels alongside the absence of both Gt and Kr. This accounts for the tight regulation of this gene's localized expression within only some three-cell-wide stripe along the A-P axis. Such regulation by multiple gap gene factors determines segmental identities for each stripy expression of the pair-rule gene network.

This is the sequential network of *Drosophila* developmental genetics facilitating early embryo patterning. The process begins with the maternal genes with transcripts initially deposited either evenly (*caudal, hunchback*) or polarized (*nanos, bicoid*) in the egg. Following fertilization, these maternal genes form early protein gradients throughout the embryo. When the MZT happens, they activate the gap genes. The gap genes code for transcription factors, via the toggle switch cross-regulations, excluding each other from their domains of expression. Having established into broad but defined zones, they modulate the expression of the next tier: the pair-rule genes. Pair-rule gene network is skipped over every other segment identity (specifying either the even or the odd numbered ones). As germ band extension takes place, the parasegments defined by pair-rule gene expression pattern into the final segmented body plan near-simultaneously. The segmentation genes are activated at this stage to demarcate the boundaries of compartmentalization. At this stage, the segments are separate from each other and cells cannot pass between established segments. Both the gap and pair-rule genes specify certain domains along the body axis as well by activating selector genes. While the formation of segments is a specific task guided by the **segmentation genes** (*e.g. engrailed*); the **selector genes** define and label the identities of each segment. The cells are **selecting** their end fate by assessing the expression of the latter set of genes. We will delve into the selector genes in more detail in the next chapter.

FURTHER READING

The linear arrangement of six sex-linked factors in Drosophila as shown by their mode of association. A.H. Sturtevant, **Journal of Experimental Zoology,** 1913. DOI: 10.1002/jez.1400140104.

The gap gene network. J. Jaeger, **Cell. Mol. Life Sci.**, 2011. DOI: 10.1007/s00018-010-0536-y.

Chapter 12: SELECTOR GENES, BODY PLANS

"SOMOS NUESTRA MEMORIA,
SOMOS ESE QUIMÉRICO MUSEO DE FORMAS INCONSTANTES,
ESE MONTÓN DE ESPEJOS ROTOS."

WE ARE OUR MEMORY,
WE ARE THAT CHIMERICAL MUSEUM OF INCONSTANT FORMS,
THAT HEAP OF BROKEN MIRRORS.

– Jorge Luis Borges, 1969.

Some of the most sensational work on selector genes, which are also called **homeotic transformation genes** by fruit fly developmental biologists, came from perturbing gene expression in the **imaginal disc** of fruit flies. The research group of Ed Lewis published seminal work on the selector genes of fruit flies in 1983. The **halteres** are diminutive winglike structures important for the fly's balance just beneath the wings. While **halteres** arise from the haltere discs in the T3 segment, **wings** arise from the wing discs in the T2 segment. Lewis's lab was able to fully convert the halteres into an additional set of wings by mutating three genes within the *ultrabithorax* (*Ubx*) locus. The compounding mutant alleles in the *Ubx* locus converts the T3 segment identity into the T2, allowing for an additional set of wings to form and replace the halteres: a fly with four wings!

Imaginal Discs

An important feature during insect development are imaginal discs. These tissues are pockets of specialized epithelia (beginning as some 20 – 40 cells) that fate commit during larval development as little stamps on certain locations along the body segments. For instance, the wing disc is present in the T2 segment of *Drosophila melanogaster* and continues to grow and expand as a patterned single sheet until metamorphosis. As the larva undergoes metamorphosis, the two-dimensional structure balloons, converting into a three-dimensional form. Now, the anterior and posterior halves of the sheet structure becomes a side of the wing. Many of the fruit fly body organs, eyes, legs, antennae, begin with a sheet-like structure that later gets sculpted into 3D shapes growing out during metamorphosis like pieces of origami.

Later on, it was realized that the Ubx locus was not the only genetic component that could change segment identities. Other researchers, including more from Lewis lab, identified more homeotic transformation genes comprising the Hox cluster (**Figure 12.1**).

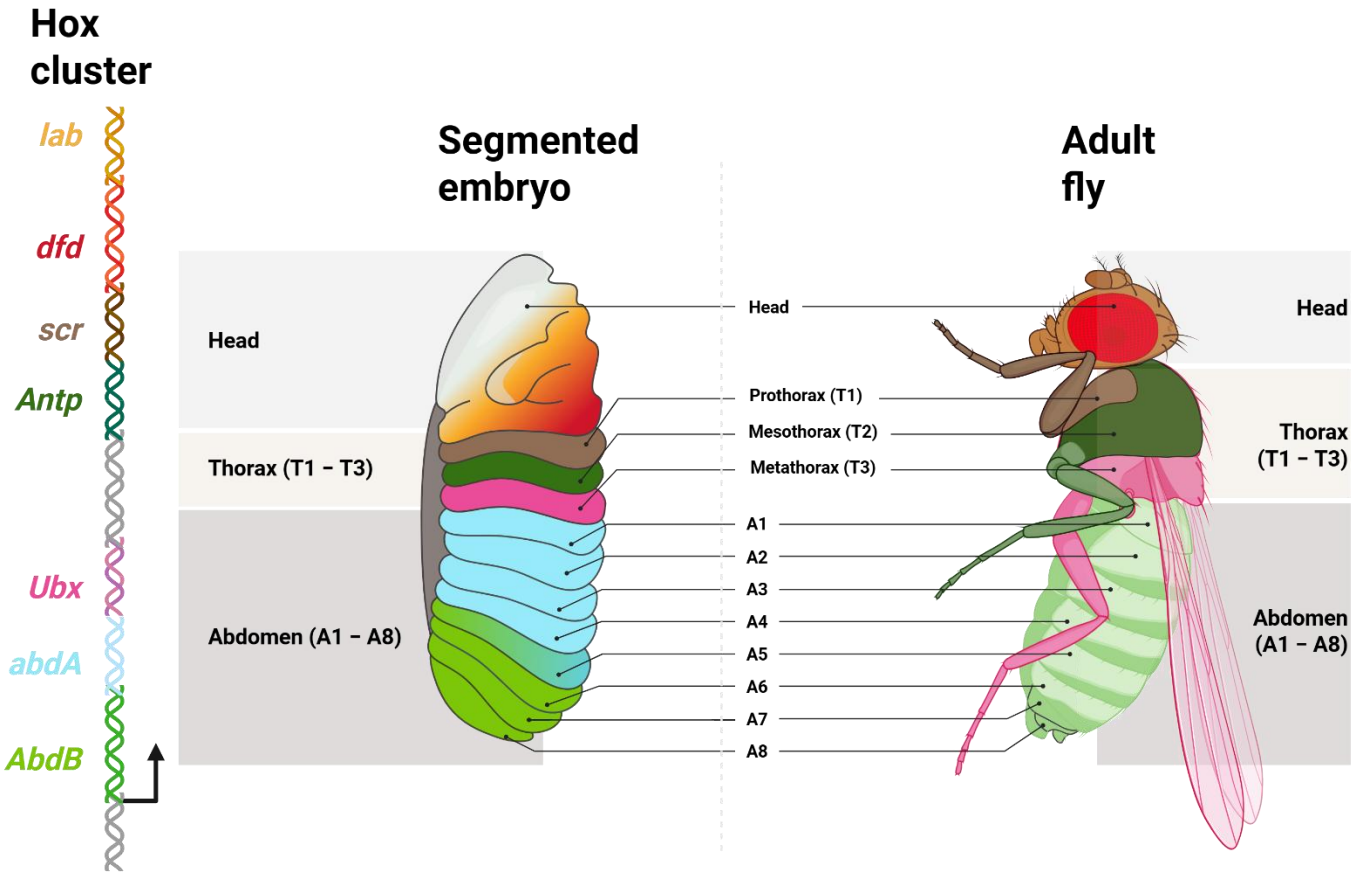

Figure 12.1 | Hox cluster in chromosome with 7 hox genes in Drosophila (left, 5′ end, bottom). Each gene is expressed in specific domains along the segmented embryo (middle, anterior top) which translate into adult body parts (right).

Let us go through in greater detail, the three loci that form the **thoracic complex** (*Ultrabithorax, abdominal-A,* and *Abdominal-B*). If all three are missing from the larvae, all the posterior segments besides the final A9 segment, become identical repetitions of T2. If only *Ubx* and *abdominal-A* are present, then segments A5 – A8 are identical to the A4 segment. In the absence of *Ubx*, T3 and A1 are replaced by T2 segments, explaining why the fly develops an additional set of wings. We see that the selector genes function combinatorially and in accordance with their domains of expression. We can imagine that the more posterior segments and differential fates an organism needs, the more of these "selector" genes will be necessary.

If these selector genes are misexpressed in different positions, they will result in homeotic transformations in phenotype, hence, they are known as **Hox genes (homeobox transcription factors)**. The Hox proteins are expressed in different domains of the larvae in an anterior to posterior order, notably following the physical order of the *hox* genes in the genomic locus. The Hox loci that encode the antennapedia (*lab, dfd, scr,* and *Antp*) complex and the bithorax complex (*Ubx, abdA* and *AbdB*) are found on the same chromosome, with the antennapedia complex being encoded closer to the 3′ prime end of the chromosome. The more anteriorly expressed a gene is, the more downstream it is located in the *hox* locus. We saw that the removal of *Ubx* from the fruit fly caused the development of additional wing structures, but it is not only the absence of *hox* genes that can cause aberrant development. Antennapedia (Antp) is worthwhile a deep dive! On T2 segment of wildtype fly larva, Antennapedia confers identity on the imaginal discs for the second pair of legs. Development of legs on T1 is due to expression of a different Hox gene, *scr* (*sex combs reduced*). However, if Antennapedia is ectopically expressed in the head region (where it is not normally expressed), this converts the whole antennal imaginal disc to a leg disc and leads to the development of an ectopic leg. As an aside, an interesting feature of arthropods is that they can grow legs pretty much from any body structure with the appropriate (or engineered) genetic cues.

The Hox genes encode a homeobox, a small (60 amino acid) DNA binding, domain that enables transcriptional control and act like a genetic timer. Tight molecular controls (DNA packaging, methylation, and other epigenetic factors) mediate their sequential expression along the chromosome from the 3′ toward the

5′ in a timely fashion. In an oversimplistic timeline, we can explain this "temporal collinearity" as, the genes situated at the 3' end of the Hox cluster (which marks anterior structures) are transcribed first, triggering a sequential activation cascade down the chromosome. Once a *hox* gene is turned on at a correct position, its expression needs to be maintained from that early developmental point into the adulthood. Essentially, the labels need to stay on, so the cells do not forget about their place; to ensure this, many epigenetic factors play a role. As we brought up in Chapter 10, the anterior cells are more advanced than the posterior during embryogenesis. Now we know a reason for that: their Hox genes are turned on earlier. And for more and more posterior identities, cells need the combinatorial effects of multiple Hox genes as they turn on.

What about other animals and Hox? The chromosome order of hox genes and their order of expression along the body axis align, *i.e.*, the temporal collinearity is valid, across species from flies to mice! The organization of loci and the zones of expression are flexible. Using mice as a mammalian example, four hox clusters (a, b, c, and d) exist in separate chromosomal loci (Hoxa1 in the anterior to Hoxa13 in the posterior, Hoxb1 to Hoxb13, and so on) instead of a single cluster in fruit flies. Hoxc6 was the first mammalian hox gene cloned by Eddy de Robertis Lab at UCLA. The reason there are additional clusters has to do with two **whole genome duplication events**. Over the course of evolution, animals in certain lineages may duplicate their genomes in seldom events providing their progeny with greater selective flexibility to evolve new genes or modulate gene expression. In the absence of selective pressure, many duplicated genes would go lost into pseudogenes. Let us compare mouse with a different vertebrate, the zebrafish. Ray-finned fish lineage underwent an *additional* whole genome duplication event long after the ancestral split of the mammals, and we observe seven hox clusters in zebrafish. Collectively, successive expression of Hox cluster genes specify more posterior positional labels along the rostral-caudal axis from flies to fish to humans.

Broader picture in evolution

We know that the Hox system is conserved across the animal kingdom. By comparing the hox genes between fruit flies and other animals, such as mice, we can get some information about the plausible hox clusters which were present in the common ancestor of vast bilateral species – sea worms, millipedes, birds, fish, frogs, arthropods like flies, and mammals like humans and mice. From comparative genomic loci analysis, we can predict that this hypothetical ancestor, *Urbilateria*, should have had 7 hox genes in its cluster. This

analysis also helps reveal how gene duplication events shaped hox clusters over the course of evolution; including the duplicated genes that are maintained, and the ones get lost between different evolutionary lineages.

As we brought up the concept of segment identities, you might be thinking back to a similar concept in previous chapters: **somite identities**. Let us look back on the transplantation experiment where a piece of PSM tissue was transplanted from a more advanced chick embryo into the PSM of a less mature embryo. In the donor, forming somites would specify the thoracic region while the host was merely segmenting cervical somites. The transplantation results in the chick embryo developing ribs within its neck region as a result of preserved somite identity. This identity is largely marked by the expression of specific hox genes. Hence the pattern of vertebrae specializes from anterior to posterior: cervical, thoracic, lumbar, sacral, caudal. This does not mean that Hox genes are the master regulators of patterning. Patterning is driven by cellular decision making, and Hox genes are turned on along the way as the embryo develops. The Hox code provides labels to the cells for them to reference what structures they will eventually develop into. Differential expression domains of Hox genes along the body axis, let us say in between chick and mice embryos, can then enable the former to make fourteen cervical vertebrae in the neck versus only seven in the latter (as well as almost all other mammals!). On the other hand, mice would make some 30 caudal vertebrae extending the tail whereas chick would suffice with nine!

Why is it that Life has selected and maintained the hox clusters in animal development? The **modularity** of the hox genes enables the evolution of different body plans. Consider the diversity of the arthropod class which includes bees, dragonflies, scorpions, lobsters, and millipedes. By altering the zones in which certain hox genes are expressed (changing where along the A-P body axis hox genes are activated), different segments flexibly attain alternative identities. As such, some organisms might have more legs by extending the segments that activate the “leg program”, whereas others can utilize the same cells in a “novel” appendage program. This can be summarized into two major principles for the success of Hox code: **a) Metamerism:** segmented body plans enable modular control of fate commitment. **b) Genetic programs to build organs:** With such regulatory network packages, cells can make limbs or other structures based on the label provided by hox gene expression.

We can wrap briefly:

1. The structures of an embryo are defined in the anterior-posterior axis; the more posterior a structure is, the greater number of hox genes will turn on to specify that position (**posterior prevalence rule**). As a result of this, mutations to a Hox gene result in failure to establish the more posterior fate.

2. Hox proteins are expressed in nested lateral stripes in the order their genes are located within the chromosome.

3. Mislocated expression of Hox result in homeotic transformations.

4. Hox genes are organized into chromosomal clusters that underwent large-scale duplication during vertebrate evolution, creating groups of paralogous genes across different chromosomes.

5. Animal plan diversity establishes on two principles: (1) bodies are made up of repeating segments which facilitate modularity and (2) modular genetic programs are used to build organ structures.

FURTHER READING

Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. J. Deschamps, D. Duboule, **Genes & Development**, 2017. DOI: 10.1101/gad.303123.117.

Chapter 13: THE POSITIONAL INFORMATION

"CELLS IN EMBRYOS, LIKE SAILORS AT SEA, USE TIMERS [CLOCKS] AS WELL AS SPATIAL SIGNALS [RULERS] TO TELL THEM WHERE THEY ARE AND HOW THEY SHOULD BEHAVE."

– Julian Lewis, 2008.

In our previous chapter, we highlighted the role of Hox genes in establishing identities. While the *Hox* expression keeps the positional information "label" required for development of specific structures, the

instruction by large comes from the **morphogen gradients of cell signalling**. The diversification of body plans through the positional information encoded within a Hox gene cluster allows for new body plans, new tasks, and broadly the diversification of life forms between different species. However, the Hox genes themselves are not sufficient as instructors for the specification of the plan. We understand that instructions for development largely come from cell signalling. From our previous chapters, we know that cells can influence the fate of one another through the process of induction which is followed by commitment. As cells commit to carrying out certain tasks (biological processes), they develop into a structure or form a pattern in the tissue. Around this process, these committed cells get labeled with Hox expression. As development progresses, these Hox labels will inform the cell the positional coordinate it is found within the body. What is the instructional information by which these labels are assigned to a cell?

Morphogens

Morphogen is a diffusive, form-giving substance. Alan Turning, famed British mathematician who decrypted the enigma code during WW2, first coined this term ("morpho-" form, "-gen" giving) in his 1952 paper, "The Chemical Basis for Morphogenesis". However, the idea of a form-giving substance localized within a living organism can be dated at least back to 18th century. In mid-1700s, Charles Bonnet, a French naturalist, proposed that there are localized head-forming and tail-forming substances in certain worm species. Bonnet's work primarily focused on the *Lumbriculus* worm, which you might have frequently found as a kid playing in the dirt. If these worms are split into two pieces, each half survives and regenerates the missing head or tail structures. From Bonnet's work, we can trace the initial ideas for presence of a substance within organisms to provide "positional information" during growth and development.

1. Positional information:

What is the importance of **positional identity in development?** The more explicit concept of a "body plan" for animals came in the works of Étienne Geoffroy Saint-Hilaire as detailed in Chapter 9. Saint-Hilaire claimed that the vertebrate and invertebrate body plans are similar but flipped in the dorsoventral axis. This concept of an **inverted body plan** can be accepted as historically the first mention of a positional information aligned with an axis. More specific ideas of "positional information" arose in the mid-twentieth century from

transplantation experiments conducted on *Galleria mellonella* (wax moth). For the wax moth, the cuticle pat terns of the larval stage can be inferred by observing the direction of hair growth relative to the rib position, which forms a dark straight midline in the larva. Each cuticle line is typically straight during normal larva development.

In a series of experiments in 1966 (**Figure 13.1**), the ectoderm surrounding the rib position line was removed from the larva, flipped 180° and transplanted back either aligning with the outer rib line or asymmetrically. The rib-symmetric grafts exhibited circular cuticle formations where the transplantation took place, however maintained a straight rib line. When an asymmetric graft was carried out, the rib line shifted and formed a circular pattern disconnected from the outer line. This work showed clear evidence for a positional identity for every single cell of insect ectoderm which they act upon in pattern formation. When this tissue is flipped, the positional identity memory results in a different pattern. Following these observations, researchers proposed that **the positional identity was encoded by a concentration gradient of an unknown and diffusive substance.**

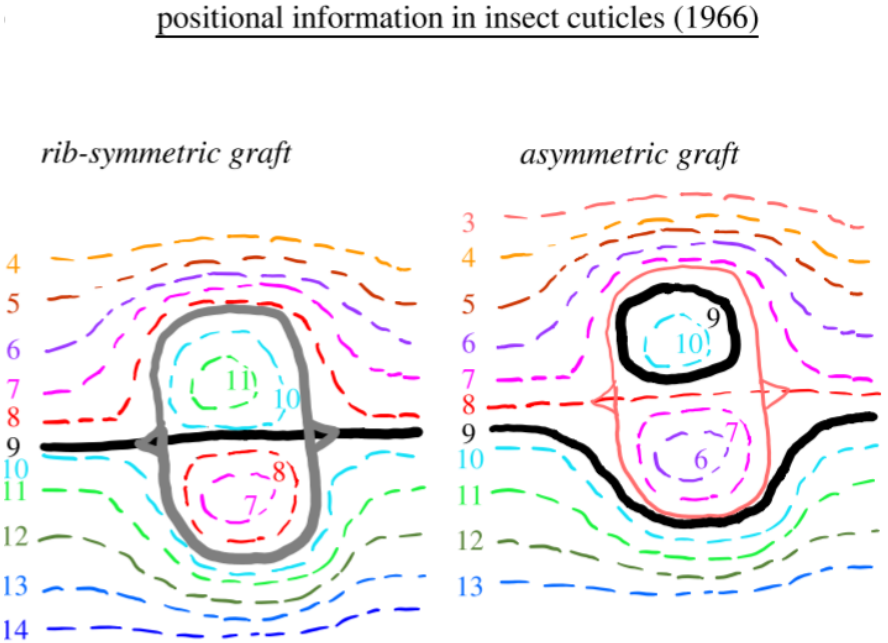

Figure 13.1 | Wax moth cuticle grafts revealing the positional information present in embryonic development. Adapted from Figure 1c, in Simsek and Özbudak, Open Biology, 2022 (DOI: 10.1098/rsob.220224).

2. Scale invariance:

The second phenomena central to the discussion is **scale invariance**, which was observed by embryologists for a long time. This concept comes from the observation that although eggs of an organism may have varying sizes, the **pattern of development is tightly conserved during embryogenesis.** Let us re-visit the early experiments of Hans Driesch, where he observed that in splitting an embryo during early cleavage into four separate masses resulted first in four smaller embryos forming, and then four small sea urchins arising. Broadly, if we think about the eggs, they are naturally variable in their size, some might be larger while others smaller, but the embryos will always be patterned the same.

A more explicit experimental investigation of this came in 1975, where Jonathan Cooke compared the growth of two *Xenopus laevis* (frog) embryos, one which was unperturbed and another which has 25% of its cells were removed from the blastula. As each developed further, Cooke observed that the blastula with lower cell mass was still able to form the full body axis and make the same number of somites when compared to the full-sized blastula. However, the somites were just "scaling" to form smaller. Altogether, the somite sizes mirror the size of the developing embryo, and the scale between the two, as well as where ear forms, where eyes form, *etc.*, remains proportionally untouched in the patterning of the organism (**scale invariance**).

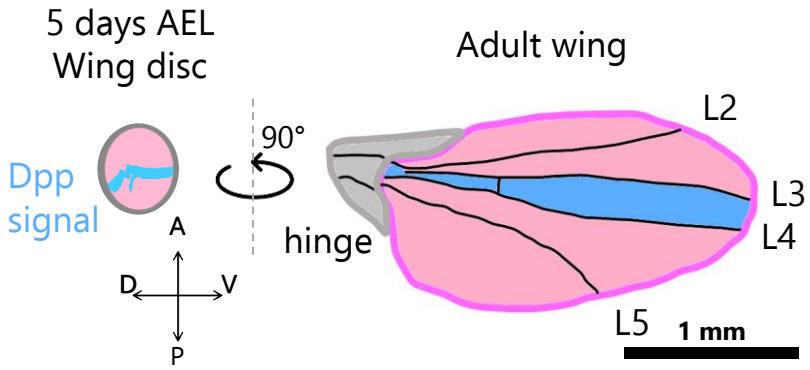

Figure 13.2 | Drosophila wing disc pouch (left) grows from 30-40 cells up to more than 10,000 cells (in 5 days after egg laying) with scaling Dpp gradient (blue) which patterns the adult wing disc A-P axis (right).

We additionally see scale invariance phenomena in pattern growth. For instance, over the course of three days, the *Drosophila* wing disc expands form some 30-40 cells into tens of thousands (**Figure 13.2**). Further

developing, this two-dimensional wing disc inflates and expands into a large three-dimensional adult organ. As the size grows, however, the anterior-posterior boundary and the structures that form, *e.g.*, location of the veins, are preserved. And we know that Dpp, the fly BMP homolog, is central to wing disc patterning and separates the anterior and posterior portions of the wing disc. The key feature here is that during growth, the initial pattern that is established is scale invariant as the tissue grows.

The French flag problem

In the mid-twentieth century, Lewis Wolpert provided a potential explanation for both **scale invariance** and **positional information**. Using the concepts brought up by Alan Turing, Wolpert proposed that if the morphogens are really diffusing within the tissues from a source located on one side of the tissue, and if the sink were on the opposite side, the tissue would end up with a linear gradient of the morphogen. Furthermore, if the tissue is growing, this linear gradient would be stretching. If cells can sense the concentration levels of the gradient, let us say at specific positions like x_1 and x_2, the relative point of these positions would change depending on the axis length of the linear gradient. Therefore, as a tissue expands, or shrinks, the pattern is preserved. This iconic concept in developmental biology is referred to as the **French flag problem.** A few years later, Francis Crick, one of the three individuals whose work led to the discovery of the DNA double-helix, produced a short manuscript with some back-of-the-envelope calculations of diffusive behavior of those hypothetical morphogens in support of Wolpert's explanation of **scale invariance** and **positional information**.

However, very shortly after its publication, Wolpert and others have noticed a very apparent problem regarding the **French flag problem** (**Figure 13.3**). In the explanation of a linear gradient, the major idea was that there was a single localized sink on one side, with the source being opposite, and that the signal would be "flowing" unperturbed over the cells. In biology, we know that cells are not passive participants! Diffusive chemical signals would interact with their receptors and often these signals are removed (through endocytosis) or degraded following receptor recognition. In this context, we see that cells are actively clearing out the morphogen tissue. This means that rather than a single distally localized sink, each individual cell in the tissue

forms a sink, which changes the clearance rate of the molecule through the tissue.

If we assume expanding the sink through the tissue, the scene changes. The rate of change being proportionate to the *amount* of a given variable means an exponential equation. Being cleared along the tissue, this means that morphogen concentration (M) will vary on the length (L) of the tissue not linearly, but

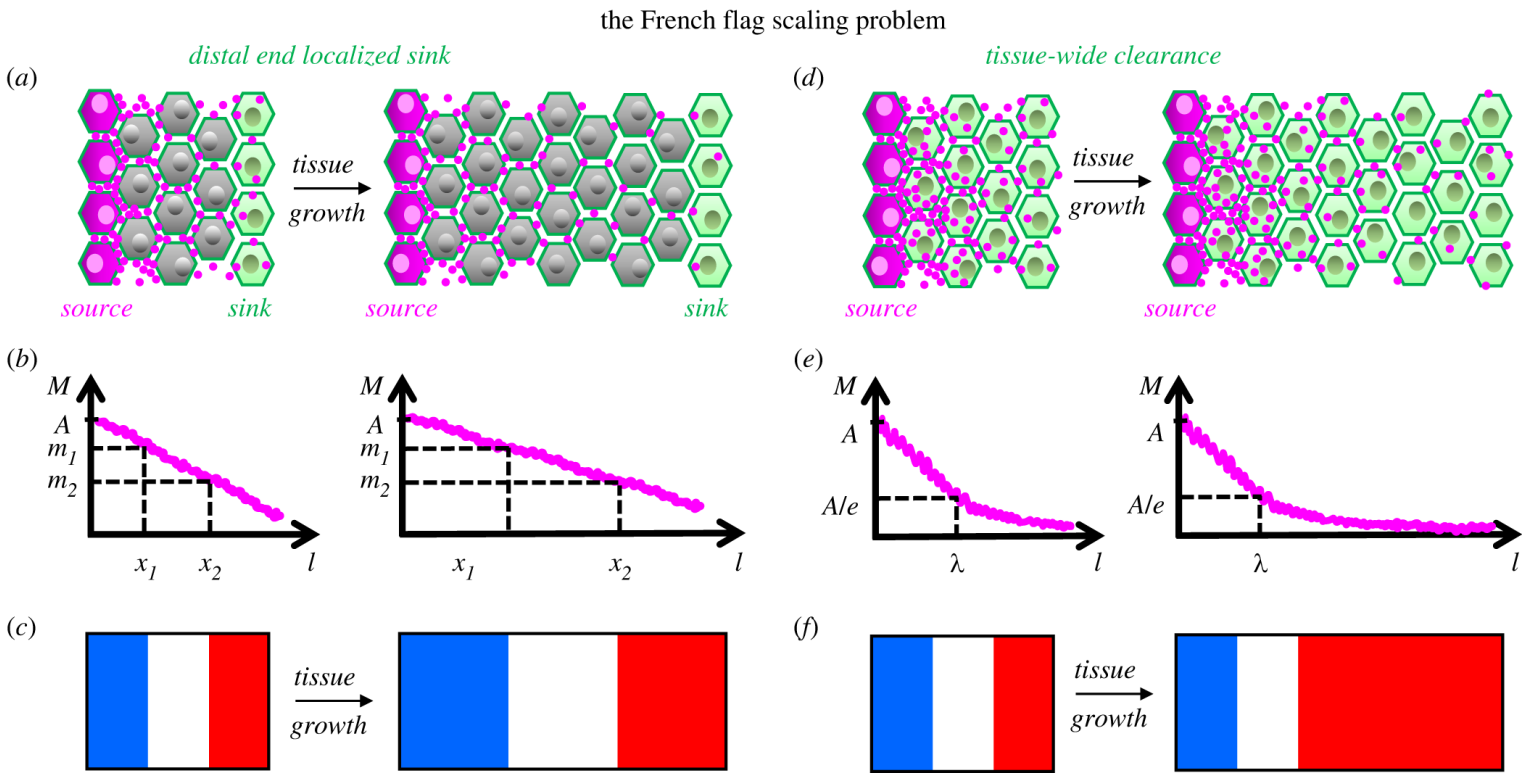

Figure 13.3 | Scaling of proportions for French flag pattern can be explained with a linear morphogen gradient (left). This requires a localized source and sink for the morphogen. If the sink expands to whole tissue, gradient becomes exponential and pattern scaling fails (right). from Figure 4, in Simsek and Özbudak, Open Biology, 2022 (DOI: 10.1098/rsob.220224).

exponentially. The trouble with this is that the tissue size cannot quite stretch or shrink an exponential equation. Mathematically, the span of exponential gradient will purely depend on two factors: The diffusion (D) and the rate of clearance (k_{clear}). Although, the French flag formulation is quite insightful to understand pattern invariance in development, linear gradients are rare to find!

With these concepts in hand, let us explore case studies of some well-studied morphogen gradients and their roles during embryogenesis.

Case Study: Three Patterning Examples by Morphogens

1. *The Bicoid gradient*

The first morphogen gradient that was molecularly discovered and characterized is the Bicoid gradient in fruit fly blastoderm. In Drosophila eggs, maternally deposited bicoid mRNA transcripts are restricted to the anterior end. In the syncytial blastoderm, where nuclei are not separated by membrane boundaries, Bcd is able to diffuse throughout the whole embryo. In this example, the source is highly localized as Wolpert proposed, originating at the anterior end, but the sinks are each nucleus. In result, researchers observed that the gradient was, expectedly, exponential when it was discovered in 1988 (Driever and Nüsslein-Volhard, Cell). Bcd is important in patterning as bicoid mutants lose all their anterior structures. More specifically, Bcd is important in positioning the cephalic furrow during gastrulation. This furrow separates the cephalic (first three) segments from the remaining 11 segments of the fruit fly.

Notably, researchers found that when they injected Bcd into varying positions of the larva it stimulated the formation of the anterior structures at that position. As we know from our previous chapter, Bcd operates through providing positional information for hunchback (Hb) expression. When the Bcd dose is doubled in a larva, the cephalic furrow shifts towards the posterior and this also expands the expression of Hb. Altogether, these experiments supported the idea of a morphogen gradient, that there is some concentration gradient which can shift the target gene expression domain. This, ultimately, can then shift the cellular pattern commitment; e.g., the formation of the cephalic furrow at certain levels of Bcd. In conclusion, Bcd as a diffusive morphogen, provide positional information for Hb expression and patterning for the head fold.

---

2. *Morphogens in neural tube patterning*

In previous chapters, we had covered importance of the notochord for inducing the neural ectoderm and patterning the neural tube. One role of notochord in this process is to act as a source for sonic hedgehog (Shh). As they diffuse, the Shh molecules interact with the neural plate and neural tube cells. With induction, the very bottom (ventral) cells of the neural tube specify as the floor plate. Induced floor plate can then act as a Shh source itself and produce further signal which expands dorsally (upwards) through the neural tube. As this is taking place, the epidermis resting on top of the forming neural tube, acts as a source for BMP signalling. Along the dorsoventral neural

tube axis, we end up with two opposing gradients, BMP and Shh; the former arising from the surface ectoderm and the latter from the floor plate. Hand-in-hand, those two gradients specify all the diverse neurons (interneurons, motor neurons, roof plate, V3, etc.) which eventually give rise to the adult spinal cord. Specific neural fate commitments depend on a cell's position along the two morphogen gradients. For instance, motor neurons require absence of BMP signal but a slightly milder exposure to Shh signal, localizing them right above the floor plate and V3 fates.

To grasp how the positional information of the neural tube arises, as a necessary and sufficient cause, we can look into the experiments by Placzek et al. (1990) where an ectopic notochord is laterally grafted in chick embryos. This resulted in an embryo with two notochords; one endogenously present beneath the neural tube, and the second one transplanted to the side. As such, the neuronal cells in proximity to the transplanted notochord begins to commit into floor plate and surrounding motor neuron fates, supporting that the diffusive Shh provides positional information. At high concentration levels, this signal can convince cells to become floor plate and motor neurons.

---

3. *Morphogens in wing patterning*

In the third and final example, we are back to fruit flies revisiting a well-studied framework for morphogen gradients: wing pattern development and wing disc growth.

*During the development of the wing disc, the fly BMP homolog, Dpp, is known to separate the anterior and posterior sides of the wing disc (**Figure 13.3**). How does this begin though? Well, it all starts with the posterior side expressing engrailed in a polarized fashion. Engrailed expression turns on Hedgehog signalling, which diffuses through the wing disc from its source: the posterior side. As the signal diffuses into the anterior portion of wing disc that lacks any Hedgehog expression, it turns on the Dpp signalling. As a side note, note that the two pathways which act in a cooperative cascade here were establishing opposing gradients in the vertebrate neural tube. In wing disc, Dpp signal begins as a very localized borderline expression, acting to separate the two halves of the wing disc. Eventually, however, Dpp as a diffusive substance also establishes a signal gradient and influence cellular proliferation and fates by activating downstream Smad transcription factors. Let us go through a few targets of Dpp signalling during wing disc development.*

As Dpp expression marks the boundary between the anterior and the posterior, activated Smad downstream additionally acts as a transcription factor that represses brinker. This repression results in high levels of Brinker expression in the anterior and posterior ends of the disc, but little to no expression in the middle. Dpp signalling also turns two other genes on: spalt and omb. Spalt expression zone is tighter and closer to the middle line of the wing disc, while Omb expression zone is broader. Notably, both Omb and Spalt are additionally repressed by Brinker. So, while Dpp represses the expression of Brinker, Brinker represses the other targets of Dpp. Altogether, this results in nested domains for Dpp signalling targets. That more or less looks like that French flag idea.

If we ectopically express Hedgehog from the anterior side of the wing, we observe that it creates a second Dpp signalling line. This results the anterior portion of the wing splitting, with two wing-like structures growing out of a single wing disc. This ectopic expression and the additional formation of a Dpp signalling line provides strong evidence that Dpp itself provides the positional information for cells to be the middle of the wing disc (position of vein 4, ***Figure 13.2****), separating the anterior and posterior compartments. Furthermore, Dpp targets Omb and Spalt specify location of other veins, e.g., 2 and 3, in the final wing pattern.*

Where is the information?

These case studies reveal the complexities of how cells can interact with morphogen gradients and expand upon the initial attempts to understand *how* morphogen gradients dissipate and the scale invariance phenomena of embryonic development. Revisiting each individual case in deeper detail of what cells respond to is helpful (**Figure 13.4**):

Case 1. We see that the early *Drosophila* A-P axis appears to respond to Bcd levels within only a specific **time window.**

Initial attempts to alter Bcd levels used genetic manipulation, for example by doubling the expression with two copies of the gene. This approach cannot capture *when* things were happening in response to the perturbation. Later on, researchers were able to change Bcd levels at specific time points and found that, rather than responding to Bcd levels throughout development, cells were only responsive to the signal at specific time windows.

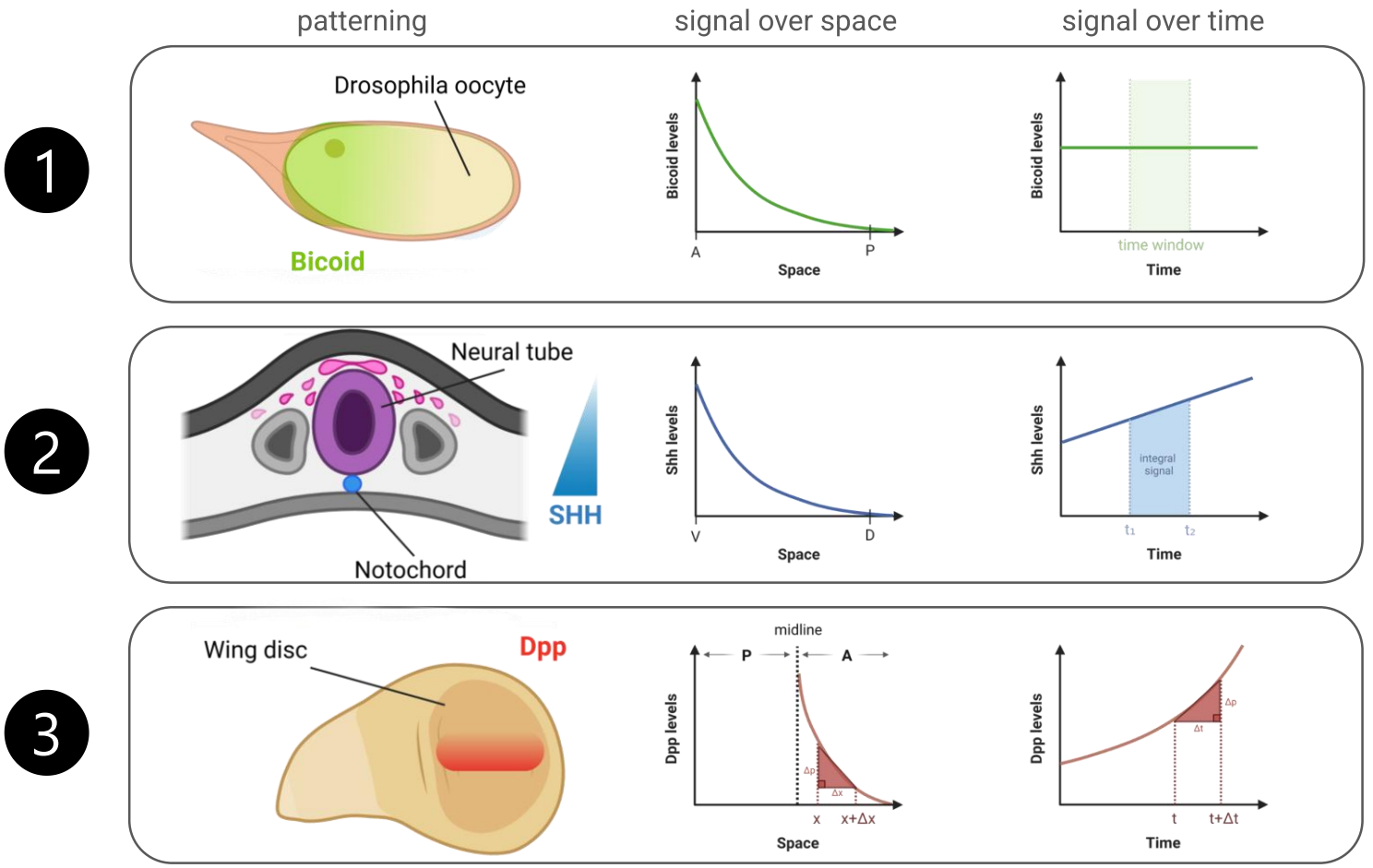

Figure 13.4 | Three classical cases for morphogen patterning and their signal dynamics in space and time (simplified): 1- Maternally deposited bicoid gradient (green) pattern the fruit fly A-P axis. 2- Sonic hedgehog (SHH) gradient (blue) sourced from notochord pattern the ventral neural fates in vertebrate neural tube. 3- Midline sourced Dpp gradient (red) in wing disc pouch pattern the A-P axis of the wing.

Case 2. Neural tube patterning appears to respond to the cumulative (**integral**) Shh signalling over time.

What was realized later on with Sonic Hedgehog signalling was that, rather than responding to the absolute **levels** of Shh, cells were instead sensitive to the **persistence** of the signal. That is, the level of incoming signal over time, rather than just the levels themselves, was more important in cellular decision-making.

Case 3. Wing disc A-P patterning appears to respond to the **fold change** of Dpp signalling.

Through the course of wing disc growth, Dpp signalling grows stronger and stronger. It was found out that the specific fold change (how quickly does the Dpp gradient change over time and/or space) is really important for cells to commit into proliferation homogeneously and maintain the initial commitment patterns as the disc grows.

Although, at the surface level, these examples support the "French Flag principle" (cells respond at a certain concentration gradient), upon further inspection and investigation we see that is not the case. In all these examples, cells are **not responding to the actual levels (a constant threshold) of signal** and instead are doing further math to respond and evaluate the signal. To sum up, while **there are morphogen gradients that cells respond to and get positional information from, it is not as simple as a line gradient with multiple thresholds.** What does allow the cell to do more math and enable complicated responses to morphogen gradients? This is largely rooted in the **effector** molecules of signalling cascades downstream of the morphogen gradients. As effector molecules network into their own and independent feedback loops, the shape of the morphogen gradient response can change from exponential to other functions, *e.g.*, sigmoidal, that guide cellular response.

Lastly, morphogen gradients can maintain their positional information content proportional with varying tissue sizes during development. However, it appears that the scale invariance with the morphogen gradients is actually transient – it only happens within a certain window rather than scaling indefinitely.

We also observe that tissue-wide morphogen clearance is causing a trouble for "scale invariance" by resulting in an exponential gradient which is hard to maintain in scale. We have seen that morphogen gradients

can theoretically provide information at multiple thresholds. That is theoretically possible and can be realized in an embryo since we observe nested expression patterns of the target genes. But what we eventually realize, as we look deeper into those biological phenomena, is that morphogen gradients do not appear to be really providing information at multiple (or many) threshold levels. Nested expression patterns arise from further gene regulatory networks at play, which we could see with the Dpp-Brinker, or Bicoid-Hunchback and their downstream regulatory networks. Likewise, a cross-regulatory network between the target transcription factors is at play in the neural tube patterning. It appears the nested expression domains, are not per se results of multiple concentration readings, rather emerge from a complicated regulatory network activated downstream.

FURTHER READING

Patterning principles of morphogen gradients. M.F. Simsek, E.M. Özbudak, **Open Biology**, 2022. DOI: 10.1098/rsob.220224.

Morphogen gradients: from generation to interpretation. K.W. Rogers, A.F. Schier, **Annual Reviews of Cell and Developmental Biology**, 2011. DOI: 10.1146/annurev-cellbio-092910-154148.

Chapter 14: QUANTITATIVE INTERPRETATION OF MORPHOGEN GRADIENTS

"ANY PROBLEM THAT IS HARD TO SOLVE BY MATHEMATICAL METHODS WILL BE IMPOSSIBLE TO SOLVE BY VERBALIZATION."

– J.M. Renier, in "The organism as an adaptive control system", 1968.

This chapter extends the discussion of morphogen gradients by focusing on the quantitative ways in which cells interpret signalling information. The central question is what a cell can compute from a signalling input in order to detect changes in signal level or spatial distribution and, in response, commit to a developmental task such as embryonic patterning or cell fate specification.

Rather than treating morphogen signalling as a purely descriptive phenomenon, we now consider how signalling networks transform extracellular information into interpretable intracellular outputs. This approach connects the earlier discussion of pattern formation to quantitative concepts such as gradient shape, slope, temporal integration, and fold-change detection.

Morphogen signalling and effector responses

In most developmental contexts, morphogen signalling occurs when a diffusible ligand binds to a receptor on the cell membrane, and the receptor transduces that information into the cell. The resulting intracellular response often involves activation of an effector molecule, which may itself function as a transcription factor or may activate other transcription factors through phosphorylation or other post-translational modifications.

Examples discussed previously include ERK as an effector in RTK/FGF signalling and phosphorylated Smad proteins as effectors in BMP or Nodal signalling. These effectors connect receptor activation to transcriptional responses in the nucleus.

This ligand–receptor model is common, but it is not universal. A notable exception is early *Drosophila* development, where localized maternal RNAs are translated into proteins that diffuse through a syncytial embryo. Because the early embryo lacks membrane boundaries between nuclei, transcription factors such as Bicoid can spread directly through the shared cytoplasm. Nevertheless, this is an exception rather than the general rule; in most tissues, membrane boundaries and receptor-mediated signalling dominate.

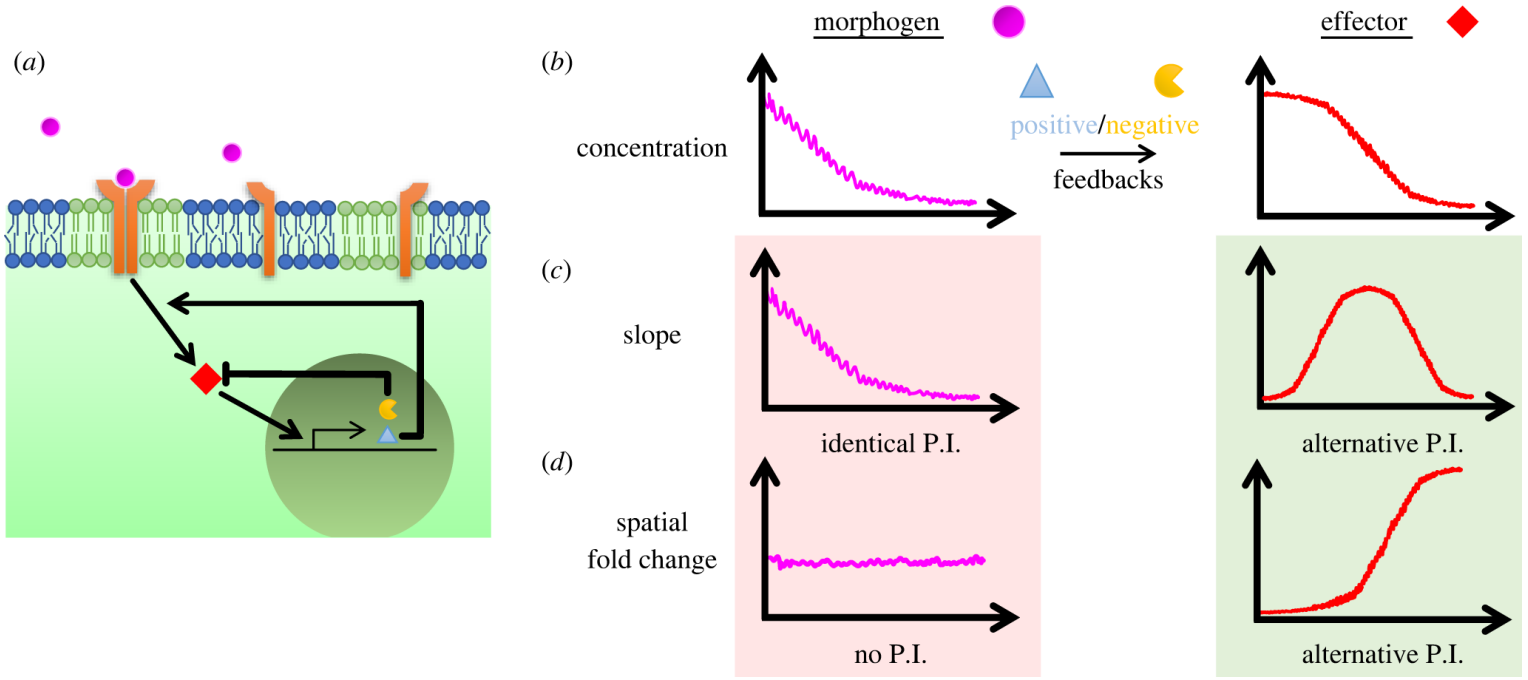

Figure 14.1 | Ligand receptor interactions, downstream cascade to the effector molecule (red diamond), and feedback converts the exponential decay gradient of the morphogen (middle) into sigmoidal shape (right). From Figure 5, in Simsek and Özbudak, Open Biology, 2022 (DOI: 10.1098/rsob.220224).

The outputs of signalling pathways can also feed back onto the pathway itself. As discussed earlier in the course, positive and negative feedback loops are common and can substantially alter the behavior of the signalling system (**Figure 14.1**).

From exponential to sigmoidal gradients

A simple diffusive ligand often forms an exponential concentration gradient across a tissue. However, once intracellular processing and feedback are taken into account, the effective signalling output can acquire a different shape. In particular, feedback can transform a simple exponential decay into a more sigmoidal profile, characterized by a plateau near the source, a steeper intermediate region, and a lower threshold region farther away.

This distinction matters because the raw ligand gradient is not necessarily the quantity that cells interpret directly. Instead, cells may respond to the downstream effector profile generated by the signalling network. Such reshaping is important in growing tissues, because a purely exponential gradient depends primarily on ligand diffusion and clearance and does not automatically scale with tissue growth. The effective output of the pathway therefore contains more interpretable positional information than the ligand concentration alone (**Figure 14.1**).

Possible cellular computations on gradients

Once the signalling output has been reshaped, cells can potentially perform several kinds of computations on it. One possibility is “slope detection”, in which a cell responds not to the absolute signal level, but to the rate of spatial change in the gradient.

For a purely exponential gradient, slope detection does not add much information, because the derivative of an exponential function is itself exponential. In other words, the shape of the information remains essentially unchanged. By contrast, if the signalling output is sigmoidal, the slope is low near the plateau, rises in the intermediate region, and falls again toward the distal end. This creates a more spatially informative signal.

A second possibility is fold-change detection in space, in which cells compare their signal level with that of neighbouring cells. For a purely exponential gradient, neighbouring ratios remain effectively constant along the tissue, so little or no positional information is gained. However, once the output becomes sigmoidal, these comparisons vary with position and can support robust cell fate decisions.

Measuring time: oscillators and delays

Before considering more complex computations, it is useful to ask how cells measure time. A single self-repressive molecule can generate a stable steady state through negative autoregulation, much like a thermostat. However, negative feedback alone does not usually produce sustained oscillations.

If one molecule activates a second molecule and that second molecule represses the first, oscillatory behavior can arise. Yet without additional features, these oscillations are often damped or overdamped. The crucial ingredient for stable oscillation is a time delay in the feedback loop. Such delays can be introduced by additional positive feedback steps, multistep reactions, or regulated degradation.

When a feedback circuit includes sufficient delay, it can produce stable limit-cycle oscillations. Once a system oscillates reliably, it provides a mechanism for time measurement: the cell can use the periodic rise and fall of a molecule as an internal clock.

Case Study: Somite Formation as a Framework for Interpreting Signals

---

*Vertebrate somite formation provides a useful framework for thinking about how cells interpret morphogen signals. In the posterior body axis, uncommitted cells in the tail region periodically form somites in synchrony with the segmentation clock. An FGF signalling gradient contributes to this process. Several hypotheses can explain how cells decide when to commit to somite formation (**Figure 14.2**):*

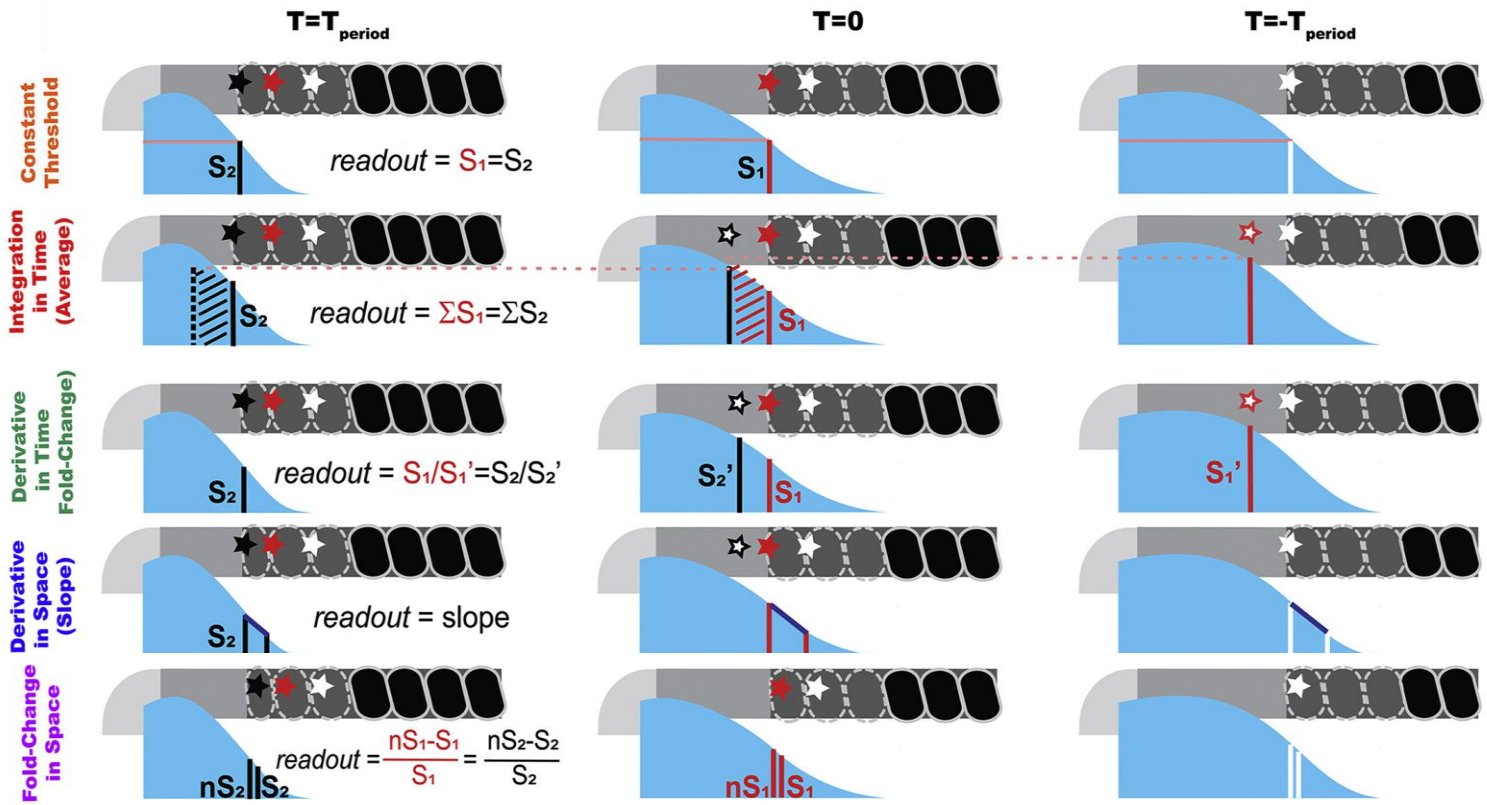

Figure 14.2 | Somites (black) segment from the presomitic mesoderm (gray, one lateral half is shown, posterior left). Specific locations (white, red, and black stars) mark the somite boundaries determined to segment in previous (right), current (middle), and next (left) clock cycle, respectively. Five hypothetical measurement scenarios are sketched in rows for comparison. From Fig. 3c in Simsek and Özbudak, Cell Reports, 2018 (DOI: 10.1016/j.celrep.2018.06.023).

1. *Threshold detection: cells respond when the gradient reaches a specific value.*
2. *Temporal integration: cells accumulate signalling over time and respond when the integrated signal exceeds a threshold.*
3. *Temporal fold-change detection: cells respond to how quickly the signal changes over time rather than to its absolute level.*
4. *Slope detection: cells respond to a particular spatial slope of the gradient.*
5. *Neighbour comparison: cells compare their signal levels with those of nearby cells.*

Distinguishing among these conceptually distinct models requires carefully controlled experiments in space and time.

Time integration in neural tube patterning

A clear example of temporal signal integration comes from neural tube patterning by Sonic hedgehog (Shh). Shh is produced ventrally by the notochord and floor plate and regulates the expression of transcription factors such as PAX7, OLIG2, and NKX2.2.

The pattern is not produced by simple activation alone. Shh activates both NKX2.2 and OLIG2, but NKX2.2 represses OLIG2. As a result, NKX2.2 becomes dominant in regions of highest Shh signalling, whereas OLIG2 is displaced slightly more dorsally. Conversely, PAX7 is excluded from ventral regions because Shh represses it, directly or indirectly within the broader network.

Experiments in culture show that different Shh levels produce different transcriptional outcomes: no Shh favours PAX7, intermediate Shh favours OLIG2, and high Shh favours NKX2.2. Importantly, when Shh signalling is strong and persistent, OLIG2 expression can initially rise but later decline as NKX2.2 accumulates and represses it.

This result indicates that cells are not responding only to the instantaneous concentration of Shh. Instead, they are integrating the signal over time. Persistent signalling allows downstream regulators to accumulate, and that accumulation changes the eventual transcriptional outcome.

Fold-change detection and molecular memory

Another possible computation is fold-change detection, in which cells respond to relative changes in signal rather than to absolute magnitude. To detect change over time, however, a cell must possess some form of memory: it must retain information about past signalling states in order to compare them with the present state.

How sense perception happens is a useful analogy. Humans often detect changes in light, sound, or smell relative to background conditions rather than in absolute terms (Weber-Fechner law in psychology). A daily life example is going nose blind, which drastically lower the response burden of living systems as they interact with their environments. Similarly, cells may respond most strongly when a signal changes substantially relative to its previous level.

A classic biological example comes from bacterial **chemotaxis** towards nutrients. Bacteria navigate chemical gradients by alternating between tumbling and longer directional runs. Their signalling network includes an adaptation module with reactions that occur on different timescales. A slower process effectively stores information from an earlier time point, whereas a faster feedback process acts on the current state. This difference in timescale creates a molecular memory.

As a result, the bacterial response is transient (doing long runs) when the signal changes and then returns toward baseline if the signal remains constant (tumbling to change direction). The system therefore responds to change, not simply to sustained level. This enables bacteria to quickly reach the source of nutrient gradient.

In eukaryotic cells, a related network motif is the incoherent feed-forward loop. In this motif, signal *X* activates target *Z* directly, but also activates an intermediate *Y* that represses *Z*. Because the direct arm is fast and the indirect repressive arm is slower, *Z* shows a transient peak when *X* changes and then relaxes back. This architecture can implement temporal fold-change detection.

Slope detection in the *Drosophila* wing disc

The *Drosophila* wing imaginal disc provides an example in which cells may respond to the slope of a morphogen gradient. In this tissue, the BMP homolog Dpp is produced in a central signalling region and controls both patterning and growth. The default state of wing disc is lateral cells being more proliferative. This proliferative state is repressed by Brinker laterally and promoted by Dpp signal centrally. Loss of Dpp signalling disrupts normal wing development. Downstream of Dpp, the pathway effector phosphorylated Mad forms a broader readout than the ligand itself. Because the wing disc grows from a small group of cells to a much larger tissue while preserving pattern, a major question is how cells decide where and when to proliferate.

To address this question, Rogulja *et al.* (2005) generated clones of cells carrying a drug-inducible, constitutively active Dpp receptor. When the drug was added, Dpp signalling could be activated locally in those clones without directly affecting neighbouring cells. The key result was that wildtype cells surrounding the high-signalling clone proliferated more strongly than expected from signal level alone. This suggested that cells were responding to the steepness of the local gradient, rather than simply to high Dpp activity.

A proposed model invokes a hypothetical molecule induced downstream of Dpp signalling. If this molecule mediates intercellular matching between adjacent cells, then neighbouring cells with similar levels could interact productively, whereas mismatched neighbours could generate a signal associated with proliferation. Although the molecular identity of this factor remains uncertain, the experiment supports the idea that cells can detect spatial differences across short distances.

Although the slope-detection model for wing disc growth is compelling, it is not sufficient on its own. Later work by Wartlick *et al.* (2011) showed that Dpp signalling levels increase over time as the wing disc grows. If the entire signalling landscape changes with time, then a simple slope-reading mechanism would fail to

explain the observed regular growth behavior. An alternative model which may explain the cellular proliferation and uniform growth in wing disc is fold-change detection in space (SFC). This way, both the cell non-autonomous response observed by Rogulja *et al.*, and the temporal increase of the signal observed by Wartlick *et al.*, may be incorporated into the solution.

This leads to an important broader conclusion: real developmental systems may combine multiple computational strategies. Cells may integrate signals over time, compare present values with past values, sense local slope, and compare themselves with their neighbours. Developmental patterning therefore depends not only on the presence of morphogens, but also on the network architectures that allow cells to compute with those signals.

In summary, a single cell is capable of surprisingly sophisticated information processing. Through feedback loops, time delays, and network motifs, cells can perform computations that allow them to interpret dynamic morphogen signals and translate them into robust developmental outcomes.

FURTHER READING

Defining network topologies that can achieve biochemical adaptation. W. Ma *et al.*, **Cell**, 2009. DOI: 10.1016/j.cell.2009.06.013.

Spatial fold change of FGF signalling encodes positional information for segmental determination in zebrafish. M.F. Simsek, E.M. Özbudak, **Cell Reports**, 2018. DOI: 10.1016/j.celrep.2018.06.023.

Chapter 15: SYMMETRY, ASYMMETRY

Within this chapter, we will focus on body symmetry, and lack thereof, in animals during development. What do we mean when we use the phrase: "symmetry"? Let us consider the jellyfish, from the Cnidaria clade. If we turn a jellyfish around its oral-aboral axis, if we roll it around, we always see the same animal shape. This rotational symmetry is referred to as **radial**. Alternatively, we could look at an animal like a human, or zebrafish, and notice another type of symmetry; that they are mirror images of themselves down the middle. This form of symmetry is referred to as **bilateral**. Many animals exhibit bilateral symmetry, including worms, snails, arthropods, vertebrates, and so on. Animals that conform to this symmetry type are part of the **Bilateria** clade.

What about some less clear examples? If we look at a starfish, we might be inclined to say it exhibits **radial** symmetry. However, this monicker does not fit exactly, although neither does **bilateral** symmetry (Unless you are talking about Patrick Star of Sponge Bob cartoon)! It is only when we look at the larva of a starfish that we can see it displays bilateral symmetry during embryonic development. Once the starfish metamorphosizes into its adult form, that symmetry is hidden. Notably, sea urchins are similar to starfish in the sense that they begin as bilateral larvae which develop into more radially symmetrical adult organisms.

If bilateral, where then is the head of starfish? Tracing the anterior posterior Hox gene expression domains into adulthood provides an interesting answer. A 2023 study by Formery *et al.* conducted exhaustive RNA sequencing in starfish which was followed by *in situ* hybridization to confirm their results and tracked the genes responsible for patterning and cell signalling. Looking at the orthologs of the Hox genes, they saw that the early expressed *Hox1* begins at the very bottom on the peripheral ends of the starfish arms, and later gene expression becomes tighter and moves upwards. Looking at what Hox genes are expressed, starfish seem to turn on the head and some trunk genes, but hardly any tail associated Hox proteins. In conclusion, the adult starfish appear pretty much as all head (with more posterior identities along the periphery of arms). Notably, although their metamorphosis converts the bilateral body axis into something closer to radial symmetry, starfish maintain the anterior-posterior order of Hox gene expression.

Maintaining outward symmetry

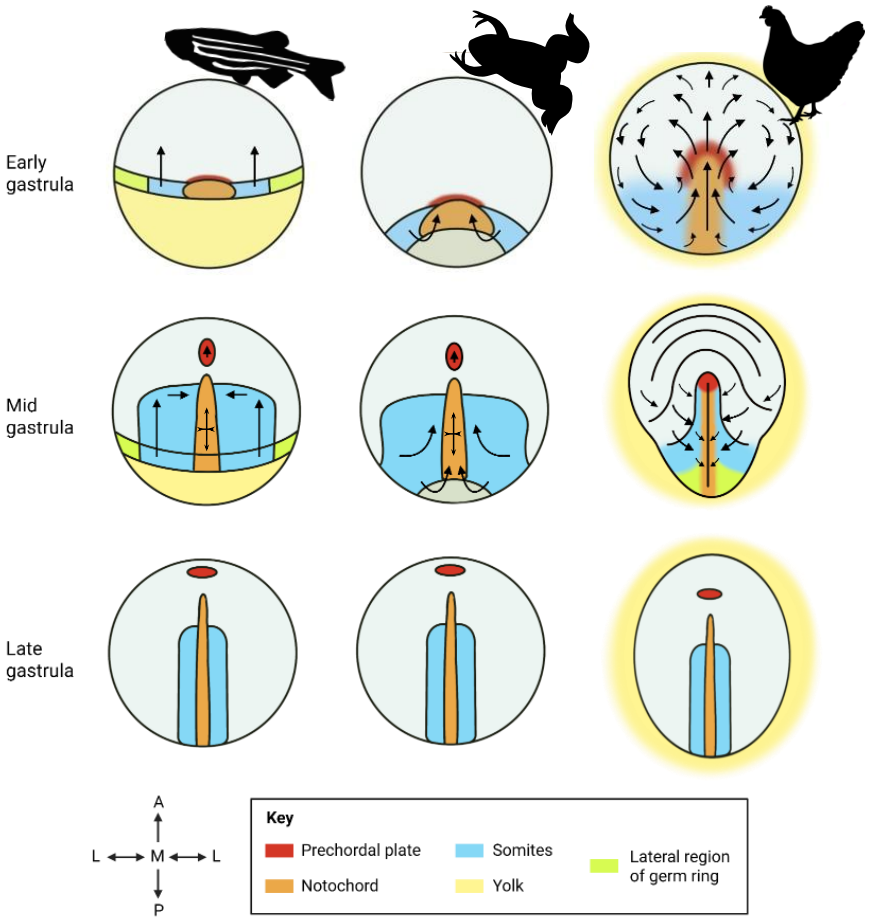

Figure 15.1 | Gastrulation movements (arrows) for various vertebrate models –fish (**left**), frog (**middle**), and chick (**right**)– put the embryos into the similar body plan (**bottom**) from different initial conditions due to yolk geometry (**top**).

The bilateral symmetry of animals comprising the Bilateria clade, from wasps, butterflies, to sharks, fish, humans, and rats, is largely driven by the **organizer**. In Chapter 9, we have outlined the role of the organizer in mediating the body plan of the embryo. The location of the organizer, which gives rise to the notochord in chordates, is situated in the middle of the gastrulating embryo and from this localization is able to split the embryo into two sides by establishing the embryonic midline between the left and right. Following the formation of the D-V and A-P axes during gastrulation, the organizer distinguishes the third (M-L) axis. How does that final axis that mediates symmetry get established?

As the organizer is situated in the midline of the gastrulating embryo, the gradients that arise from this structure (Wnt and Nodal

signalling) end up being symmetrical. The embryo will end up with a symmetrical formation as the organizer drives embryonic axis formation. This is best exemplified looking at early gastrulating, the BMP and Wnt signalling are symmetrical, and as the AP body axis forms vitamin signalling continues to be in line with the AP axis, and Fgf signalling goes from the tail to the somites. For vertebrates, which is what we will primarily focus on in this chapter for simplicity's sake, this is why bilateral animals have similar structures.

Another feature, besides the positioning of the organizer and subsequent symmetry of the gradients, driving bilateral symmetry is the symmetry of cellular movements during gastrulation (**Figure 15.1**). As convergent extension results in similar body plans between differing organisms, we see that cellular movements follow the symmetry of the organizer. In examples which does not follow convergent extension, such as the chicken development, there is still a symmetry to cellular movements arising instead from **polonaise movements** (**Figure 15.1, right**). The cells, in this context, swirl either to the left (left side) or right (right side) of the primitive streak keeping the embryo symmetrical.

The final feature maintaining bilateral symmetry in Bilateria is axis elongation. As the cells ingress into the tail portion of the embryo, they undergo a Wnt-mediated epithelial to mesenchymal transition which drives an increase in disorder and motility. As the cells remain ordered further away from the tailbud, and the Wnt signalling, the unidirectional growth of the axis resembles how icicles or stalagmites form; the bottom cells behave like a fluid and the top cells resemble solids, driving elongation. Why is this balance of mesenchymal disordered cells to epithelial ordered cells important? Researchers decided to track individual cells in the zebrafish tail bud and compared a wild-type to a cadherin-2 knock out, which would cause the resulting embryo to lose the localization of the epithelial to mesenchymal transition from the bottom of the tail bud and homogenizes the disorder in the embryo. It was observed that, in comparison to wild-type zebrafish, the cadherin-2 deficient embryo maintained symmetrical axis elongation but at a far slower rate. Increasing the disorder does not affect the symmetry of the axis development and only the rate. What about when the researchers *decreased* the disorder in the tailbud? When over-expressing the Wnt-inhibitor, *notum1a*, researchers observed that the tail began to bend and lose symmetry. With low Wnt signalling, the epithelial to mesenchymal transition cannot take place, and the cells are still gastrulating but not increasing in disorder or fluidity. The more structured push of the cells and the cell motility cause vortices of cell movement to form

on one side of the tail, causing axis bending. So, we can see there is a sweet spot balance between getting rapid axis elongation (by minimizing disorder) and maintaining symmetry (minimizing structure of the gastrulating cells).

Following gastrulation and the establishment of embryonic symmetry, organogenesis begins to take place. As vertebrates develop their lateral organs, such as limbs, which grow out of the major body axis, how do they maintain symmetry? And how is it that the growth of each limb is similar, and we do not end up with arms of differing lengths, for example. This begins with looking at the anterior-posterior patterning, which we know are driven by hox and transition genes. The patterning of the AP axis is split into three distinct groups: the head, trunk, and tail. It is at the transition between these groups that limb outgrowth takes place; where the head and trunk transition, we see forelimb, while in the trunk to tail transition we see hind limb formation. The positional information derived from the somite boundaries maintains the symmetry of where limbs grow on the left and right side of the organism. These conclusions are not universally true but hold for most vertebrates. Of course, there are examples that deviate, such as a snake, which lacks any limb outgrowth.

Looking at the limbs in greater detail, we will now answer how the rate of outgrowth remains constant. How is it that two tissues that are separate during embryonic development nevertheless come to the same size? Well, this begins with the limb bud, which is the structure that derives from the embryonic mediolateral axis. Akin to the imaginal discs from *Drosophila*, limb bud outgrowth is dependent on BMP signalling (Dpp in flies, the insect ortholog). In vertebrate limbs, BMP signalling is controlled by two signalling centres, the apical ectodermal ridge (AER), in the most distal part of the limb, and the zone of polarizing activity (ZPA) in the posterior of the limb. The AER is an Fgf signalling centre, while the ZPA is a Shh signalling centre. In this signalling feedback, the middle of the limb bud has another player, Gremlin, which is a BMP inhibitor. From these, you have both negative and positive feedback loops playing into BMP regulation. We know that Shh activates BMP, which represses Fgf, and Fgf is required for Shh activation. As Gremlin inhibits BMP, it enables Fgf signalling. The coupling of these feedback loops maintains a tight control on BMP and makes it such that the growth pace, which is referred to as **proliferation time** and tells the cells to divide for a certain duration depending on the BMP signalling, is uniform between the two limb buds. As the bud continues to grow, the ZPA and AER grow further apart, and the signalling distance becomes too great for diffusion to be effective.

This causes a break in the feedback mechanism and signals to the bud to stop growing. From this early pattern, cartilage, bone, and muscle formation follow. The tight control of BMP signalling ultimately ensure that the right limb went through the same number of division events to the left limb. If you would have even one extra round of division, the organism will have one limb that is twice as long as the other! To have the limbs be the same size on each size, the number of cell divisions needs to be the same.

Establishing internal asymmetry

Although vertebrates exhibit outward bilateral symmetry, we know that our internal organs are not arranged that way. For instance, the liver and heart are not localized to the midline. Why is it important for those organs to be arranged asymmetrically? Two asymmetrical organization can be stated for inner organs: *situs solitus*, which is normal anatomy (what most people have), or *situs inversus*, which results in a mirror opposite orientation. The total reversal of organ arrangement, **situs inversus totalis**, is a rare event happening in one of every 30, 000 births. However, this alternative orientation is not detrimental to function, and someone might go their whole life not knowing their organ arrangement is different unless clarified with medical imaging. It is only when **situs ambiguous** or **heterotaxy**, which is when the organ arrangement asymmetry is lost and the heart is positioned in the midline, takes place there are consequences. Heterotaxy results in a loss of internal structure and disrupts the communication taking place between our organs. Around 90% of neonatal births with heterotaxy possess heart defects requiring surgery, and with corrective surgery around a quarter would die from the procedure. The misarrangement of internal organ asymmetry is serious and detrimental to healthy development.

So, how is it that cells break the established left-right symmetry we just outlined? In the universe, there are no macroscopic features that distinguish an object on the basis of it being "left" facing or "right" facing. However, in the molecular world, it is something observed more readily. We are all familiar with molecular chirality and how left-handed biomolecules are common in life versus right-handed counterparts. In development, the molecular asymmetry becomes translated to the embryonic scale by the **left-right organizer** (LRO). This transient organ, which links gastrulation to somitogenesis, have different names depending on the organism (Kupffer's vesicle for zebrafish, Hensen's Node in chick at the end of the primitive

streak, Node in mice, Gastrocoel roof plate in frogs). LRO is located in the embryonic tail, where the notochord and the neural tube meet posteriorly, a location also referred to as the **chordoneural hinge**, as a small void space, *i.e.*, typically filled with fluid. This fluid is surrounded by motile ciliated cells, which are the source of the symmetry breaking that drives functional organ arrangement.

How is it so? Motile cilia possess cylindrical axonemes which are built by a 9+2 doublet structure of their microtubules. A central pair of microtubules sit within nine outer doublet pairs connected by nexin links which form the cylindrical structure; radial spokes connect the outer and central microtubules doublets. For each outer doublet, there are dynein molecules bound. The outer dynein attaches to the shell of the cylinder, while the inner dynein motor faces towards the central microtubule pair. The dynein motors walk unidirectionally on the axoneme to keep the cilia beating; consuming substantial amounts of ATP (~10, 000 molecules every minute!). While there are many models attempting to explain how this takes place, what we know with better certainty is that the directional movement of the motors causes an elastic force that accumulates and causes dyneins to drag on the microtubules. As a result of this stress, the cilia beat unidirectionally.

It is this molecular movement which breaks the macroscopic symmetry by triggering two major changes. First, signalling gradient molecules will be sent towards the left side with the unidirectional movements of the cilia. Secondly, as the cilia beat along the same orientation, the fluid flow is higher to the left side promoting calcium transients. These calcium transients expanding into the left side of the embryo as waves drive the degradation of *Cerberus* RNA, which acts as an inhibitor of Nodal signalling. Consequently, this causes an asymmetrical expression of Nodal favouring the left side of the embryo. And, as we learned in previous chapters, Nodal expression activates the expression of more Nodal, eventually diffusing the signal from the left side to the lateral plate mesoderm and the endoderm. In parallel, Nodal drives the expression of Lefty, its diffusive repressor, which ensures the expanding Nodal signal centre doesn't evade its venue of action on the left side. These processes consequentially lead to asymmetrical patterning of the heart, gut, and other internal organs.

In 2023, the journal of *Science* published a study by Djenoune *et al.* that provided direct evidence that motile cilia are responsible for embryonic symmetry breaking (at least in zebrafish!). The research group used a custom combined microscopy to place the left right organizer, with the ciliated cells, through a light sheet

path allowing them to image the structure in three-dimensions. In the meantime, a third objective was used to introduce an optical trap, which is heavily focused light that forms an electric field gradient. This electric force "traps" and controls the movement of structures within the focused beam. After treating the embryos with morpholino such that ciliated movements in LRO were blocked, they used the optical trap to capture the stationary cilia. By tuning oscillations of the optical trap, the researchers were able to selectively return motility to individually captured cilium. Researchers longitudinally followed those manipulated embryos, measuring Cerberus expression at early somite stages (8-10) and later observing where the heart forms. By forcing a single cilium to beat in a specific direction with the optical trap, scientists were able to restore left-sided Nodal signalling and observe functional and properly placed heart formation.

FURTHER READING

Molecular evidence of anteroposterior patterning in adult echinoderms. L. Formery *et al.*, **Nature**, 2023. DOI : 10.1038/s41586-023-06669-2.

Patterned disordered cell motion ensures vertebral column symmetry. Das *et al.*, **Developmental Cell**, 2017. DOI : 10.1016/j.devcel.2017.06.020.

Cilia function as calcium-mediated mechanosensors that instruct left-right asymmetry. L. Djenoune *et al.*, **Science**, 2023. DOI: 10.1126/science.abq7317.

Chapter 16: ORGANOGENESIS I: LIMB DEVELOPMENT

"ALL MODELS ARE WRONG, BUT SOME ARE USEFUL."

– George E. P. Box (British statistician, 1976).

As the embryo develops, the globe-shaped limb bud structures erupting bilaterally at the head-trunk and trunk-tail intersections become patterned into well-shaped limbs. The rate of patterning is species dependent. This process takes place in humans within some mere 15 days! We shall note that a significant embryonic growth also takes place during this period of patterning. As we delve into limb patterning, we will keep in mind that limb patterning is a topic which not about patterning alone, but it is patterning *and* growth. For the remainder of this chapter, we will focus on the limb development of tetrapod animals, primarily human and chicken.

The limb bud extrudes from the lateral plate mesoderm, which is different than the paraxial mesoderm we previously discussed. The paraxial mesoderm, which is positioned to the left and right of the notochord and the neural tube, gives rise to the somites. The lateral plate mesoderm, on the other hand, is distinct and lies next to the paraxial mesoderm – many inner organs arise from the lateral plate, for example, the heart which develops from the anterior lateral plate that is closer to head. The limb buds emerge as 3D structures with AP and DV axis, as well as proximo-distal (PD) axis as they grow further away from the embryonic body.

The development of the limb bud is controlled by two signalling centres, the zone of polarizing activity (ZPA) and the apical ectodermal ridge (AER). The ZPA is located in the posterior end of the limb bud and polarizes the bud along the anterior-posterior axis, while the AER is in the ectoderm and is located by the middle of the limb bud, at the distal central ridge, polarizing the PD axis. These signalling centres are key to forming the three major Cartesian coordinates which define the limb bud: the AP, which forms the thumb to the pinky fingers, the DV, which forms the back of the hand to the palm, and the proximal-distal (PD) which forms the base of the arm to the end of the fingers.

Harrison's limb bud grafts

The experimental evidence supporting the existence of Cartesian coordinates in limb buds dates back quite a while ago. In 1925, Ross Harrison, a researcher from Yale University, had conducted a series of transplantation experiments on the limb buds from salamander embryos. Early in development, when the limb bud was extruding from the lateral mesoderm, he extirpated the limb bud mesoderm tissue of one embryo and transplanted it into another. In these transplantations, he either maintained position of the limb bud (orthotrophic) or placed them somewhere else along the body axis (heterotropic). Besides changing the positioning of the limb bud on the body axis, he also would flip the tissue on its anterior-posterior axis, dorsal-ventral axis, or both (8 combinations), changing its orientation in the new embryo. In the hundreds of transplantations he conducted, he was able to report on approximately 60 successful transplantations, where the engrafted mesoderm integrated into the embryo. From these experiments, he was able to understand that the limb bud has polarity in all three coordinates.

Molecular markers of limb bud outgrowth

Following the molecular biology revolution of the 1980s, we were able to identify the proteins and genes that are important for limb bud growth from the embryonic axis. It was discovered that the lateral plate mesoderm, where the limb buds extrude from, gets positional information from the somites enabling the medial structures to know where they are and where to grow the limbs. While there are many complicated signalling cascades, transcription factors, and signalling molecules, for the purposes of this course we will focus on two T-box transcription factors: Tbx4 and Tbx5. Both of these transcription factors turn on Fgf10 signalling, which is a key signalling molecule important for the initial outgrowth of the limbs. If we were to look at the expression levels of Fgf10 in the chick embryo, we would see that this signalling factor is well localized to two positions in the AP that specify to the forelimb and hindlimb. Additionally, it has been shown that if you engraft transgenic cells, secreting the Fgf10 ligand, into another position of the chick lateral plate mesoderm it drives growth of additional limb formation at that location. These grafting experiments reveal that Fgf10 is sufficient for limb bud initiation. However, as researchers tried to identify the localization of the Fgf10 activators Tbx4 and Tbx5, they found that the expression levels of these factors differed from Fgf10. Tbx5 expression is strictly isolated to the forelimb, while Tbx4 is strictly isolated to the hindlimb. At that point, the grafting experiments

become more exciting! Cells can be grafted in between forelimb and hindlimb, where an intersection pool of Tbx4 and Tbx5 expressing cells are present. The consequence of this mixed expression results in the formation of an extra limb that had both wing-like and leg-like structures, being feathered in the anterior and clawed in the posterior. Tbx4/5 expression encodes the differential identity of forelimbs and hindlimbs.

The expression of Fgf10 eventually drives the activation of Fgf8 in the ectodermal signalling centre, AER. As the Fgf8 signal diffuses from the ectoderm into the mesoderm, it reinforces Fg10 expression, creating a positive feedback loop. The cross regulation of Fgf8 and Fgf10 allows for sustained limb bud outgrowth. During this developmental process, Fgf8 expression is tightly localized to the AER and controls proximal-distal axis growth.

The apicoectodermal ridge: AER

The understanding that limb bud outgrowth and patterning are shared processes during limb development arise from a series of chick embryo transplantations conducted on the AER. First, the excision of the AER from a developing chick embryo halts limb outgrowth, confirming that the AER must be present for development to take place. Conversely, when additional AER tissue is inserted into the ectoderm surrounding the forelimb bud, one can observe that the whole wing is duplicated. Effectively, this results in two distinct PD axes forming, driving wing duplication. Altogether, these results indicate that loss of the AER stops wing development, and that excess AER signal drives not only growth but the formation of a PD axis.

Further experiments assessed the importance of the mesodermal mesenchymal cells in limb bud growth and patterning. Again, using chick embryos, researchers extracted mesenchymal cells from the hindlimb bud and transplanted them into the forelimb bud. What they ultimately observed was that the wing began to grow the distal structures of the legs, such as fingers and claws. This supports that the limb bud cells, with their differential Tbx4/5 expression are already committed to making their respective limb structures. Outgrowth continues because the AER is intact, but the patterning changes and resembles the legs. What about if non-limb bud mesenchymal cells are transplanted? In this case, researchers observed that **no** outgrowth takes place, despite the presence of the AER. From these results, we can infer that although the AER is important for

outgrowth, the mesenchymal cells must also be **competent** for the formation of the limb, which requires expression of either Tbx4 or Tbx5.

Lastly, following the excision of the AER, researchers replaced this signalling centre with Fgf8 soaked beads and observed limb outgrowth. What they found was that they could restore outgrowth and wing patterning to wild-type levels. Although the Fgf8 soaked beads resulted in minor patterning differences, these observations were sufficient to conclude that Fgf8 is **necessary and sufficient** for outgrowth, and the importance of AER in developments comes from its release of this signalling molecule.

As the forelimb develops, our arms as humans and wings in birds such as chicks, there are three primary structures we see being formed during outgrowth: the stylopod, the zeugopod, and the autopod. The **stylopod** forms the first structures between the shoulders and elbows, while the **zeugopod** forms the forearm to the wrist and the **autopod** forms the hands and digits. When assessing the consequences of AER removal from the limb bud, researchers observed that arrest of outgrowth was dependent on *when* excision takes place. If the AER is excised earlier, while there is only a single cartilage mass forming, then growth is arrested following formation of the stylopod. However, if you excise a bit later, after the stylopod has split into two cartilage pieces, the zeugopod forms. Excising after that results in the formation of everything to the wrist, leaving a forearm structure that lacks the autopod formation. These observations reveal that as outgrowth takes place, there are different structural identities forming along the proximal distal axis.

There are many models that attempt to explain this coupling of growth and patterning, and we will focus on two that are strongly supported by experimental evidence. The first, the **opposing gradients model**, suggests that AER signalling is coupling with retinoid acid (RA) signalling, which we know takes place at the proximal end of the limb bud, opposite to the AER signalling centre. This model posits that if both RA signalling **and** Fgf signalling are high, this would activate the formation of the stylopod. However, as outgrowth continues, this pushes the signals away from each other. As each gradient gets lower, medium levels of both would signal for zeugopod formation. As the limb continues to extend and grow, ultimately the distal end would experience high Fgf signalling, from the AER, and low RA signalling, driving commitment to autopod formation.

The other model is the **progress zone (timer) model**, which implicates highly prolific mesoderm cells found underneath the AER. As Fgf is a growth factor, the high signalling concentration from this centre drives the formation of highly proliferative cells, **progress zone cells**. This model instead posits that these progress zone cells possess an intrinsic timer, and that early in the timer these cells would be making the stylopod. As the timer continues, the cells would then progress to forming the zeugopod, and when the timer runs out the cells commit to making the autopod structures.

The first model suggests that both RA and Fgf signalling gradients are equally important for the growth and patterning of the limb, while the latter implicates FGF's role as a growth stimulating factor alone. Both models, although not in agreement with one another, are supported with experimental evidence. It could be that both could be accurate or wrong, but there are no clear answers, yet!

Zone of polarizing activity: ZPA

We can next bring up the second signalling centre, ZPA, which is the Shh signalling centre in posterior-distal zone. Shh signalling in limb bud development patterns the perpendicular axis, the anterior and posterior. What does pattern in this context mean? If we think about both the stylopod and the zeugopod, these independent structures are more or less symmetrical and most of the asymmetry in the limb is in the autopod (consider your thumbs and the rest of your fingers). We know that ZPA is important in digit identity from key experiments conducted in chick embryos. When a second ZPA is transplanted into a limb bud in the anterior, a mirror image of the digits is forming. In chickens, they have three notable digits 2, 3, and 4 (the first digit is very small and often ignored) and with an extra ZPA signalling centre, we see a 4-3-2-2-3-4 pattern is emerging rather than the classical 2-3-4 along the A-P axis.

These observations can also be related to the **polydactyly** as a developmental birth defect, which occurs when an offspring has more than five fingers. This can range from having a single extra digit or all fingers being mirror duplicated. Polydactyly is famously observed in cats as well, famously named "Hemingway cats" because the author was gifted a six-toed white cat in the 1930s. With the advent of genomics, researchers were able to identify a causal molecular link to polydactyly in humans. Rather than being due to mutations in the Shh signalling pathway directly, instead a distal enhancer region named ZPA regulatory sequence, ZRS,

exhibit mutations in individuals born with polydactyly. This is an impressive case where the altered expression levels rather than direct nonfunctional mutations of a developmental regulatory gene are causing the digit phenotypes.

Having discussed the Shh and FGF signal centres, we should emphasize that a third signal centre essential for limb patterning is BMP which is high in the central mesodermal cell of the growing bud. Cross regulation between FGF-Shh-BMP pathways through Gremlin and others is essential to maintain a uniform proliferation and growth succession as discussed in the previous chapter.

Separation of digits

Another important feature in the formation of the limbs is programmed cell death, **apoptosis**. The digits separate from one another as the tissues between them undergo apoptosis– the digits themselves are protected from apoptosis. This process takes place both in the forelimb (the fingers) and the hind limb (the toes). The mesoderm BMP signalling pathway is what instigates the programmed cell death, and we know now that apoptosis is only controlled by the mesoderm cells and is not dependent on the ectoderm. This was experimentally shown by assessing differences in digit separation between ducks and chickens. As we know, chicken hindlimbs end with pronounced digits and claw-structures, while ducks harbor webbing between their toes. Switching the ectoderm between the duck and chicken by removing the sleeve-like ectoderm and transplanting it on the other species' bud mesoderm, scientists observed that it was mainly the mesoderm identity that determined digit separation. When the duck ectoderm was placed on the chicken mesoderm, the foot did not display significant webbing. Additionally, when the chicken ectoderm was placed on the duck's mesoderm, the limb developed more similarly to a regular duck with pronounced webbing.

Palm reading the fate of cells

The final axis that defines limb bud is the dorsal ventral (DV) axis, which is dependent on Wnt signalling. On the dorsal ectodermal side of the limb, we know that Wnt signalling is taking place and driving the expression of the Lmx1b transcription factor in the mesoderm. As this is taking place, the mesodermal BMP signalling is driving the expression of *Engrailed* (*En1*) in the ventral ectoderm. The ectodermal expression of *En1*, driven by the mesoderm BMP signalling, and the mesodermal expression of *Lmx1b*, driven by ectodermal

Wnt signalling, hand in hand establishes the polarized DV axis. From these signalling gradients, we see that ectodermal identity derives from mesodermal BMP signal, while mesodermal identity is defined by ectodermal Wnt signal. The polarized expressions of Lmx1b and En1, which are two transcription factors, polarizes gene regulation along the DV axis. Lmxb1 mutant mice grow footpads on both sides of their feet, unlike the wild-type developing footpads in the ventral alone. In this genetic background, mice do not make dorsal structures in their autopods, supporting the importance of these transcription factors in axis formation.

Case Study: Thalidomide

During the late 1950s and early 1960s, a German pharmaceutical company tested and formulated the drug, thalidomide, which was used to treat morning sickness during early pregnancy. The company conducted a series of tests on cell culture, in animals, but did not assess their drug in pregnant animals. The extent of their tests led them to believe that thalidomide was safe for consumption, and it was placed on the market. Following its use, it likely caused, as a lower ballpark estimate, tens of thousands of babies to have developmental birth defects, most notably arrested limb development. Although many countries banned its use, others continued to allow the product to be in the market, and it resulted in a medical disaster. Alongside halting limb development, babies had developmental defects in their heart, eyes, and ears.

Why the drug resulted in these defects is not sharp clear, despite nearly five decades of research took place to understand the etiology. A key observation found is that the drug seems to only drive developmental defects when used in a certain timeslot. If used after those few weeks range, it does not harm embryonic development. One of the strongest models supported by experimental evidence is that, in this window of being detrimental to

development, thalidomide prevents proliferation in progress zone cells. Although the cells are exposed to Fgf signalling from the AER, they are unable to divide and this eventually stunts the limb development, where the shoulders immediately commit to autopod formation. Persistent, elevated levels of Fgf signalling seem to drive this precocious commitment to autopod fates. One thing we can note is that these observations seem unsupportive of the opposing gradient model we have previously discussed. As proliferation is stunted, these cells would be exposed to both elevated levels of RA and Fgf signalling, yet nevertheless limbs develop to autopods and not stylopods.

FURTHER READING

Establishing the pattern of the vertebrate limb. C. McQueen, M. Towers, **Development**, 2020. DOI: 10.1242/dev.177956.

Chapter 17: ORGANOGENESIS II: HEART DEVELOPMENT

In this chapter, we will cover, more briefly, the development of the heart and circulatory system during embryogenesis. We will first focus on the sequence of events and place less emphasis on the molecular details of how cellular decisions are made for the moment. The heart development has been studied since the times of Aristotle (350 B.C.), the genetic elements that lead to heart defects is a busy area of research. However, we have somewhat limited quantitative understanding of the cellular processes.

Heart development is something mesmerizing! One of the best memories for people, including myself, once they began doing embryology, especially with an accessible model organism like zebrafish, is looking through the microscope and watching the embryo's first heart beats.

Let us begin with the mise-en-scéne for the heart in the embryo. The heart arises from the lateral plate mesoderm (LPM) – the same germ layer structure that the forelimbs and the hindlimbs come from. More specifically, heart arises from the anterior LPM. It is important we set the scene by describing the relative position of the lateral plate mesoderm to other embryonic structures; as their signalling crosstalk heavily influences organogenesis and later development. In the midline of the early embryo, we know that the **neural tube** rests above (dorsal) the cylindrical **notochord**. These two structures are flanked by **the paraxial mesoderm**, which segment into somites. Adjacent to the paraxial mesoderm is the **lateral mesoderm**, and in the space between these two structures is the **intermediate mesoderm**. Although we did not have time or space to introduce the intermediate mesoderm in greater detail so far, it is good to know that the **gonads** and **kidneys** form from that.

The heart is formed by many complex cell types and structures, including the cardiomyocytes, which are the muscle cells that keep the heart beating; the endocardium and the epicardium, the latter of which gives rise to the arteries that feed the heart. Another important structure of the heart is Purkinje fibres which are critical in keeping the heart beating by setting the beating frequency. All these complex and diverse cell types arise from the same **progenitor cell** population present in the lateral plate mesoderm.

Heart from the lateral plate mesoderm

In the lateral plate mesoderm, the heart arises from a specific **progenitor cell** population that are initially evenly distributed throughout. Although this cellular population begins uniform between the distal-midline (proximodistal or mediolateral) axis, they will begin to migrate towards the midline as the lateral mesoderm plate becomes polarized from *lefty* expression. As these progenitor cells come together, they form a crescent structure on the midline of the embryo referred to as **the cardiac crescent**. One of the proteins important in this migration process is Foxp4; in mouse studies, mutations to *foxp4* results in the formation of two progenitor cell populations, to the left and right of the embryo, which cannot coalesce together in the midline.

There are many genetic disorders or underlying causes that might cause migration to fail. Regardless, stunted migration gives rise to **cardia bifida**, which can result in the formation of two heart structures rather than just one. Neonates with this condition are unlikely to survive this developmental abnormality because the circulatory system will not run properly. More broadly, cardiac defects are the most prevalent birth defects, occurring at a frequency of 1 in every 100 livebirths. Although as medical technology advances some of these can be addressed surgically, heart malformations are still attributed to 5-10% of all newborn deaths. In large part, this prevalence is due to the complexity of the heart and the many steps that it takes to form it properly.

Following the formation of the cardiac crescent two fields of cells will form: the **first heart field** and the **second heart field**. The first heart field acts like a scaffold for the formation of the tube and cardiac progenitors to fuse for the midline. As we discuss the formation of heart tube, it is helpful to "compare and contrast" with another tube formation event of early embryogenesis: the neural tube. The neural sheet folds into tubular structure with the help of oriented cell divisions flanking laterally. Those divisions push the lips of cell sheet towards the midline, bending and fusing together to make the tube. With the heart, the case is a bit different. The heart is making a tube likewise from a planar structure of cells, but those cells are not a continuum of epithelia that can be pushed with oriented cell divisions. Rather, it is active cell migration that drives tube formation for the heart. As the progenitor cells migrate towards the midline, this movement causes them to push against each other and results in the formation of the tube from the first heart field cells. Then, the second heart field comes into the play and shapes the tube by adding cells to its anterior and posterior ends. As this happens, the heart tube will develop an arterial pole at the anterior, and a venous pole at the posterior. Cellular

tissue mechanics and blood flow will cause the heart to loop in 3D and further structures such as the outflow tract, and the ventricles form through this **looping process**. Hand in hand with this process is chamber formation where cells also induce each other to make valves and connections between the heart chambers. Broadly, we can conclude that the first heart field is important for the formation of the tube, while the second heart field helps form the loops.

Key cell signalling events of heart formation

As every other development process during embryogenesis, heart formation depends on cell signalling. The signals received by the anterior lateral plate mesoderm starts the process in a very location specific fashion. During early development, the neural tube, above the embryonic midline, is secreting Wnt signalling. This Wnt signal diffuses to the lateral plate mesoderm; but to form the **heart plate**, Wnt must be repressed. How does this repression take place? Underneath the lateral plate mesoderm, the anterior portion of the endoderm asymmetrically produced Wnt inhibitors including Dkk and Cerberus, diffusing these signals to the mesoderm. These inhibitors block the Wnt signalling released by the neural tube to impact anterior LPM. With the repressive signals from the anterior endoderm, the heart progenitors begin forming the heart plate. On the other hand, high Wnt signal leads to vascularisation, angiogenesis, and hematopoiesis in the posterior LPM which is not experiencing Wnt inhibitors. These cells will form blood vessels (vascularisation and angiogenesis) or blood cells (hematopoiesis). An additional signal then comes into play.

BMP is required both for the heart and circulatory system fate. BMP is inhibited in the embryonic midline, deterring the formation of heart or blood cells directly in the midline. BMP inhibitors like Chordin and Noggin are released by the notochord, which protect the midline from developing into any heart or circulatory structures. Altogether, the anterior endoderm, also referred to the **pharyngeal endoderm**, both represses Wnt signalling and provides BMP ligands to drive heart formation, while the notochord secretion of BMP inhibitors creates two populations of progenitors around the midline of the embryo. Heart formation is driven by low Wnt signalling and high BMP signalling, and additionally, Fgf signalling which is triggered by BMP and has an intricate feedback balance in between. In the absence of any Fgf signalling, essential proteins that are required for heart muscle contractions to take place (*e.g.*, connexins) are not produced. For BMP signalling, we

know that phosphorylation of Smad changes gene expression. The two we will focus on in this chapter are Nkx2.5 and Mesp1.

Key transcription factors

Nkx2.5 is a transcription factor and a very early marker for cardiac fate commitment. In fruit flies, Nkx2.5 homolog is referred to as *tinman* (from Wizard of the Oz), and flies with mutations to this locus do not develop hearts. Notably, in the fly, *tinman* is commonly expressed throughout the mesoderm early in gastrulation, but following germ band expansion, its expression is limited to the dorsal mesoderm. At the end of gastrulation, expression is limited to the dorsal closure where the heart eventually arises from. Let us put the side note that fly hearts do not have chambers, instead are cylindrical tubes, and return back to amniotic heart development.

Mesp1 is also a transcription factor, but rather than driving the expression of heart-specific proteins that enable the heart to beat, Mesp1 prevents the reversal of cardiac fate commitment in cells. Mesp1 does this by repressing the expression of numerous transcription factors, including Foxa2, Sox17, Goosecoid, and Brachyury, and prevents the heart field cells from reverting to notochord, endoderm, or paraxial mesoderm fates. By repressing these TFs, Mesp1 ensures that commitment to becoming part of the heart is maintained. Additionally, Mesp1 turns on two other T-box transcription factors, Tbx5 and Tbx20. In previous chapters, we covered how Tbx5 plays a role in the formation of the forelimb, in this context it specifies between the left and the right ventricles; high Tbx5 drives the left ventricle, while high Tbx20 drives the right ventricle to form. This polarization is maintained by Tbx5 inhibiting Tbx20 expression, and vice versa (toggle switch motif). So, as we have the separation of left and right ventricles, we will focus on a fifth transcription factor, Foxh1. Foxh1 is activated by Mesp1, and in turn, reinforces Mesp1 expression in a positive feedback loop. Foxh1 is important for the patterning of the anterior heart field, which includes the atrium. Alongside all this, Mesp1 is also important to promote cell migration of the progenitor cells to the midline of the embryo.

Looking at it from a more structured scenario, we can see that the first heart field gives rise to mostly the left ventricle and that it is important for the formation of the tube. But, eventually, all the cells of the first heart field only contribute to the left ventricle, and all other heart components are built by the second heart field cells. These second heart field cells split through their expression of Nkx2.5, Mesp1, and the final transcription

factor we will cover in this chapter, Tbx1a. Tbx1a is essential for all the derivatives of the second heart field to properly pattern: the atrium, the right side of the ventricle, and some non-heart tissues including facial, jaw, and airway structures.

Heart looping and septation

Following the activation of all the signalling networks, cellular commitments, and early structures formed by the first and second heart fields; the heart undergoes **looping** for further mechanical changes. Although the embryo has formed the heart tube, the heart requires more pattern and structure for maturation. The looping and chamber formation process happen hand in hand, between the Day 15 and ~Day 50 during human development. By Day 21, the second heart field cells have been added to the anterior and posterior ends of the tube, and looping begins taking place. Proper looping depends on the left-right asymmetry of the embryo caused by *lefty* expression. As the tube continues to loop, it curves and causes the arterial progenitors to come together at the top. But, as this is taking place, the heart also requires septa and valve formation to take place at the same time. Septation of the heart is largely controlled by Tbx5, which is localized to the left ventricle in amniotes.

Septation (formation of distinct heart chambers) varies between animals. In a fruit fly, their hearts are a single tube, while fish have two chambers, amphibians and lizards have three, and amniotes and birds have four. Why are there these differences? A Nature study in 2009 by Koshiba-Takeuchi and colleagues revealed that differences in Tbx5 expression primarily drives this observation. Researchers looked at Tbx5 expression levels throughout the course of heart development in frogs and lizards (3-chambers), turtles (almost 4-chambers), and chicken, mice, and humans (4-chambers). In frogs and lizards, Tbx5 was expressed throughout the ventricles, and these expression levels were sustained throughout development. However, in turtles, over the course of development, Tbx5 expression becomes more localized. Finally, for the four-chambered animals, scientists found that Tbx5 expression is highly localized to the left ventricle, enabling septa formation. Altogether, these experiments revealed localized Tbx5 expression drives septa formation. Notably, in humans born with ventricular septal defects (VSD), which is caused by an absence of Tbx5 expression (**Holt-Oram Syndrome**), they also fail to form septa and develop 3-chambered (lizard-like phenotype) hearts.

The circulatory system

Now that we have covered formation of the heart; we will shift to a brief overview of how the circulatory system is formed. The process of forming blood vessels and blood cells is a separate developmental process from heart formation. As we previously mentioned, if Wnt levels remain high in the lateral plate mesoderm the cells will form vessels and blood cells. Three important terms for this process are:

1. Vasculogenesis: differentiation of vascular endothelial cells *de novo* from mesoderm.
2. Angiogenesis: formation of new blood vessels by sprouting or remodelling of existing blood vessels which refers to further patterning and pruning to form the circulatory system.
3. Hematopoesis: formation of blood cells.

All three vascular structures form from cells committed to the circulatory fate, which are also referred to as **blood islands**. These **blood islands** give rise to both the vasculature and the blood stem cells.

Let us focus on the angiogenesis part a little further, which is the maturation of the vascular system. Initially what we have is cells that either undergo vein or artery fate commitments, and this cellular decision is dependent on Notch signalling. If Notch signalling is high, the cells commit to forming arteries, but if Notch is low, they commit to forming veins. Notch, via lateral inhibition, enables homogenous tissue to split into two different fates and keeps a balance between the formation of veins and arteries during development.

The final concept we will cover is **angiogenic sprouting**, which is required for the formation of capillaries. Sprouting is dependent on two cell populations, **tip cells** and **stalk cells**, mostly taking place between somites (Intersomitic vessels). Sprouting itself requires a single tip cell to lead the event. In the lead (tip) cell, there is a high amount of vascular endothelial growth factor (VEGF) signal, and Notch signalling also inhibits other cells from initiating sprouting through lateral inhibition. The tip cell, therefore, must not only be high in VEGF signal, but also in Notch. As Notch is repressed in the neighbouring cells, they become stalk cells and follow the leading tip cell. Once sprouting takes place, the sprouts must find each other to connect and form capillaries. That process is referred to as pruning, remodeling, and maturation.

FURTHER READING

The deployment of cell lineages that form the mammalian heart. S.M. Meilhac, M.E. Buckingham, **Nature Reviews Cardiology**, 2018. DOI: 10.1038/s41569-018-0086-9.

Reptilian heart development and the molecular basis of cardiac chamber evolution. K. Koshiba-Takeuchi, *et al.*, **Nature**, 2009. DOI: 10.1038/nature08324.

Chapter 18: ORGANOGENESIS III: EYE DEVELOPMENT

The eye is a window for animals connecting organism's information processing unit, the brain, with the surrounding environment. Over the course of life's history on Earth, those complicated structures evolved convergently at many lineages, mostly coming in two structural flavours. We can begin first with the one most familiar for us, the vertebrate eyes, building on a lens-camera system. In that structure, we start with a lens which focuses the light on the back focal plane inside the eye. The lens is covered by a transparent and protective cornea which also contributes to focusing and a circular muscle, iris, adjusting the amount of light falling on the lens. Then, at the back of the eye, the retina function like a camera sensor. Retinal neurons, as specialized photoreceptor cells, are catching the focused image falling in and converting this sensation into electrical signal communicated to the brain through the optic nerves criss-crossing through embryonic midline at the optic chiasm. In many flies, we observe a quite different eye system, the compound eye. The compound eye is formed of a packed array of ommatidia, individual vision units each sporting small lenses and photoreceptors, together functioning like an array of small eyes processing the image from different angles and sensitivities.

Puzzled with the complexity of the eye with parts in need of each other to be functional, Charles Darwin, noted: "To suppose that the eye with all its inimitable contrivances… could have been formed by natural selection, seems, I confess, absurd in the highest degree". We will revisit this note near the end of this chapter as a case study.

Development of eye

The development of vertebrate eye is a classic example of induction. In the head region, the neural ectoderm gives rise to the brain, and within the forebrain a tissue called the optic vesicle forms. This folded neural epithelial tissue induces the overlying surface ectoderm to form the lens placode, which later invaginates and forms the lens.

How do we know this? If the optic vesicle is removed, no lens forms, so the information required for lens formation comes from the optic vesicle. Tissue other than the optic vesicle, if implanted, does not induce lens

formation. Likewise, an optic vesicle cannot induce a lens if it is grafted to a more posterior location. In other words, the responding cells must be competent to receive the signal and become lens placode and then lens vesicle. These experiments show that signalling across tissues is required for proper eye formation.

Another point we know from earlier lectures concerns specification and determination. Very early in gastrulation, for example in frog embryos, there is already a region of cells that corresponds to the presumptive eye region. We know this because if those cells are labeled, they later contribute to the eye. The interesting point comes from transplantation experiments. If cells from the presumptive eye region are taken from an early gastrula and transplanted into a host embryo at a slightly later neural stage, into a region that would normally contribute to somites, those transplanted cells can adopt somitic fate. By contrast, if cells are taken from a later stage and transplanted to the same location, they still form eye tissue rather than somites.

What does this show? It shows that the eye program begins somewhere between these two stages. At the earlier stage, the cells are associated with future eye formation, but they are not fully committed. At the later stage, the eye program has already been activated strongly enough that the cells continue to form eye tissue regardless of where they are placed. Thus, the cells are specified early but become determined later.

We can observe this process *in situ* in electron microscopy images of chick embryos. At first, the optic vesicle is not yet apparent. When the optic vesicle forms, it causes the surface ectoderm above to thicken. This transiently thickened epithelial tissue is called the lens placode. The optic vesicle then folds into two epithelial layers, which we begin to call the optic cup. At the same time, the lens starts to fold inward from the surface ectoderm, forming the lens vesicle, and later a fully formed lens appears. In the retina, the inner layer becomes the neural retina, while the outer layer becomes the retinal pigmented epithelium, which becomes heavily pigmented and therefore appears dark in the embryo.

At the level of signalling, the optic vesicle is part of the forebrain. In the forebrain there are several signalling centres. At the very anterior neural ridge, near the tip of the head, FGF signalling is high. In the dorsal and especially the cortical midline region, Wnt signalling is high. In the more ventral region, Sonic hedgehog signalling is high. These three signalling gradients—FGF, Wnt, and Sonic hedgehog—work together to specify the eye fields and form the bilateral optic vesicles. This is essentially the anterior neural tube being carved into forebrain structures, including the eye.

The sensory placodes

How is the presumptive eye tissue specified? Signalling story of eye development is a coin with two sides: one for lens and other for retina. During early gastrulation, we first observe cranial (head) placodes (transient thickenings of tissue) in the ectoderm surrounding anterior prechordal (endoderm) and neural (ectoderm) plates. Lens placode forms at the outskirts of neural crest cells encircling the neural plate alongside the olfactory (nose, in anterior) and the otic (ear, in posterior) placodes. In Chapter 9, we had discussed that neural ectoderm gets specified through inhibition of BMP signalling while BMP high regions specify as epidermis. The middle ground between neural plate and epidermis has two extra fate choices which neither become part of epidermis nor the neural tube. Where sustained Wnt signal give rise to transient BMP signal, neural crest cells are specified. If the Wnt signal is inhibited and transient and the transiently activated BMP signal takes over, then the presumptive placodes are determined. Lastly, following this intricate balance of Wnt and BMP inhibitions, the Shh signal in the midline provided by the underlying prechordal plate comes into play. As the Shh signalling sharply splits the embryonic midline (recall the cyclopia phenotype discussed in Chapter 3), both the placodes and the eye field split into two presumptive regions.

Placodes also undergo characteristic morphogenetic behaviors. As discussed earlier in the course, microtubules can elongate epithelial cells, and elongated epithelial cells often invaginate. This is similar to neural tube folding. Thus, placodes begin as thickened epithelia and then either invaginate as tissues or undergo delamination, in which cells leave the epithelium. Depending on the placode, one or the other behavior predominates. In the end, placodes give rise to different sensory or neurosecretory cell fates. Different placodes give rise to different structures. The olfactory placodes form the sensory epithelia of the nose. The otic placodes give rise to the sensory cells of the inner ear. The lens placode gives rise to the lens. In this sense, placodes are crucial for establishing the organism's connection to the outside world.

The lens placode is interesting for two reasons. First, unlike the other placodes, it does not give rise to neuronal tissue. It simply forms the lens. The retina, by contrast, comes from the neural tissue of the optic vesicle. Second, the early placodal crescent appears to have broad lens-forming potential. We know this because markers such as Eya1, Pax6, and Six3/Six6 are expressed there.

Master regulators of eye induction

Eya1, Pax6, and Six transcription factors are deeply conserved. *Eya* stands for *eyes absent*, *Pax6* corresponds to *eyeless*, and Six genes are homologous to *sine oculis*, meaning "without eyes." All three names reflect eye-defective phenotypes. To keep things simple, we can focus our attention on Pax6.

Pax6 mutant animals do not develop functional eyes, from flies to mice (*eyeless* in Drosophila, and *Small eye* in mice). From frogs to humans, reduced PAX6 dosage cause **aniridia**, a defect in which the iris does not form properly. More severe loss of PAX6 function is not compatible with normal embryonic development. In homozygous Pax6 mutant mice, the embryo is missing normal eye formation and also shows other defects in head morphology. In fruit flies, mutation of the *eyeless* similarly prevents eye formation.

The role of Pax6 is so deeply conserved through evolution is that when *M. musculus Pax6* gene is artificially expressed in fruit fly antenna, normally antenna-forming cells commit into making an ectopic compound eye in that position instead. Conversely, fly *eyeless* gene can functionally rescue eye development in mouse *Pax6* mutants. Similar observations also come from other organisms. For example, Pax6 enhancer activity can drive GFP expression specifically in the eyes of planaria and fruit flies. Those experiments show how deeply conserved the eye developmental program is across evolution and highlight Pax6 as a necessary and sufficient transcription factor for eye induction.

Pax6 is not acting alone. The Eya, Pax6, and Six genes regulate one another, and there is positive feedback among them. Once one of them is activated, the others are turned on, and together they reinforce the eye program at the same place to ensure that eye development proceeds properly.

The anterior neural tube

On the second side of the coin, the eye field is specified from the anterior neural tube. This portion of neural tube where both Wnt and BMP are inhibited becomes able to turn on Otx2 transcription factor which then accumulates and eventually releases activators of *Pax6* master regulator from repression.

A set of transcription factors under Pax6 regulation (some listed earlier) then reinforces through positive feedback loops a ventral portion of the forebrain as the eye-forming field. This field will give rise to optic

vesicle which will induce the surface ectoderm into lens placode through secreted factors *e.g.*, Fgf8 and BMP4 signalling. Lens placode will also signal back to the optic vesicle with specific FGF secretion. This mutual induction cascade will trigger invagination of the optic vesicle as optic cup and cellular expansion and separation of lens placode from the surface ectoderm as lens vesicle. Two layers of the optic cup cells further differentiate into neural retina and pigmented epithelium, expressing *Vsx2* and *Mitf* gene markers, respectively. Likewise, lens vesicle further differentiates into primary and secondary lens fibres, packed with transparent crystallin proteins. Thus, all the parts fit together. This returns us to the point raised at the beginning of the chapter: the parts of the eye do not usually arise independently in development.

Lens vesicle is induced to express *Sox2* (through BMP activation from the optic vesicle) besides the *Pax6* master-regulator gene. Sox2 and Pax6 transcription factors bind cooperatively to the enhancer sequence of δ-Crystallin gene facilitated by consecutive binding motifs for each transcription factor. Expression of crystallin initiates the lens differentiation while the FGF8 secreted from the optic vesicle also maintains the persistence of crystallin expression until the lens fibres fully differentiate.

As the optic vesicle had induced overlying ectoderm to become the lens vesicle, shortly after the lens vesicle cells separate from the surface ectoderm, they stimulate their overlying ectoderm into cornea fate. Suppression of Wnt signalling is key for this induction. Corneal induction later recruits neural crest cells to the site to furnish the cornea with corneal-specific extracellular matrix. Secondary lens fibers and the corneal tissue which together focus the incoming light unto retina are under continuous repair throughout the lifetime of the animals. New lens cells keep being added as external layers from the lens equator. Corneal epithelium displays continuous turn-over and regeneration to fight off accumulation of reactive oxygen species.

The optic cup gives rise to neural and pigmented retina epithelium layers in the back of the eye as well as the muscular iris tissue where it touches the lens vesicle to control the size of pupil (light input). In the neural retina, the photoreceptor cells absorb photons and convert them into electrical signals. Broadly, there are two types: rods and cones. Rods are highly sensitive to low light but do not distinguish wavelengths well. Cones are important for colour vision but require brighter light. That is why colour discrimination is poor in dim light.

The retina

The neural retina has three layers and contains several cell types. Retina neuroblasts are competent to differentiate into different fates; including photoreceptor (light-sensitive rod, colour-sensitive cone) cells, ganglion cells, bipolar interneurons, and notably Müller glia. Among the neuronal types, the most important to remember are the photoreceptors, which lie next to the retinal pigmented epithelium; the bipolar cells, which relay signals through the retina; and the ganglion cells, which send the output to the brain. Müller glia cells support the tissue and – as discussed in connection with Notch signalling in Chapter 2 – the choice between glial and neuronal fates depend on lateral inhibition mechanism resulting in a salt-and-petter type dispersion of those cells.

Among the eye types most familiar to us is the vertebrate eye. One notable feature of the vertebrate eye is that light enters from the front, passes through several layers of retinal cells, and only then reaches the photoreceptors. Other lineages have vastly different solutions, such as the compound eye of insects or the cephalopod eye of octopus and squid. Vertebrate and cephalopod eyes look superficially similar, but they differ in internal organization. In the octopus' eye, connecting neurons lie behind the photoreceptors relative to incoming light. Why is it so in vertebrates?

Photoreceptors are under intense metabolic demand because light detection involves signal amplification in organized membrane structures and therefore generates considerable oxidative and metabolic stress. Photoreceptors will not be quite long-lived unless they are continuously replenished by oxygen supply and gets cleaned off whenever they die off. This continuous support, replenishment, and waste handling task is carried out by the retinal pigmented epithelium and underlying choroid where the blood vessels connect to the eye. If we had the photoreceptor cells on the other end right next to the light, that'd catch more light but would lack the proper access to blood supply since it's a bad idea to mix the bloods with the nerves (Recall the blood-brain barrier!). That is how the vertebrate eye eventually ended up evolving in a functional solution.

Retinotectal mapping

The lens can focus the light, and the retina can detect it, but that is not enough. The electrical signal produced by ganglion cells must be connected correctly to the brain. While the bodies of ganglion cells are

based in retina their axons extend to the optic tectum, connecting the eye to the brain for vision. An interesting feature of the vertebrate eyes that those axons connect to the brain by crisscrossing the midline, which is the optic chasm. The crossing of nerve fibers in the optic chiasm was first documented with animal dissections by Ismail Jurjani (1042-1137), a central Asian physician who noted its role in preventing double vision (**diplopia**).

Optic chiasm is important for us to track how vision is processed in vertebrate brain (**Figure 18.1**). In a frog or fish embryo for instance, the right eye retina connects to the left optic tectum and vice versa in a topographic mapping. Dorsal side of each retina maps to lateral tectal region, ventral retina to medial tectal region, nasal (closer to nose/medial) retina to posterior tectum, and temporal retina to anterior tectum. This arrangement is important because a convex lens due to optics inverts the image on the retina. The brain's wiring preserves and interprets the spatial order while ensuring we do not perceive the world upside down. In mammalian brain, mapping further involves lateral geniculate nucleus (LGN) and visual cortex where nasal nerves from left eye combine with temporal nerves from right eye to map on right LGN and right visual cortex, and vice versa. In humans and other primates, visual cortex mappings are more complex (fine-tuned).

With differences, all vertebrates feature a single optic chiasm, and the basic topographic mapping idea is similar. The importance of retinotectal mapping is shown by experiments in frogs in which the optic nerve is severed, and the eye is rotated by 180 degrees before reconnection. After regeneration, the frog may strike in the wrong direction when trying to catch prey. This indicates that the neural mapping is highly specific.

A key signalling system involved in this mapping is the Eph/ephrin system. Eph receptors and ephrin ligands are contact-dependent guidance cues. Unlike Notch signalling, which is also juxtacrine, Eph/ephrin signalling often mediates repulsion, helping keep cell populations separate. During development, Eph/ephrin signalling is used in similar separative tasks including germ layer separation and somite boundary formation.

In the retina, Eph receptors are expressed in gradients, with higher levels in the temporal retina and lower levels in the nasal retina. In the optic tectum, ephrin ligands are also expressed in a gradient, higher posteriorly and lower anteriorly. Axons with high receptor levels are more sensitive to ephrin-mediated repulsion and therefore stop earlier, in more anterior regions. Axons with low receptor levels extend farther into posterior regions. In this way, receptor and ligand gradients together create the correct retinotectal map (**Figure 18.1**).

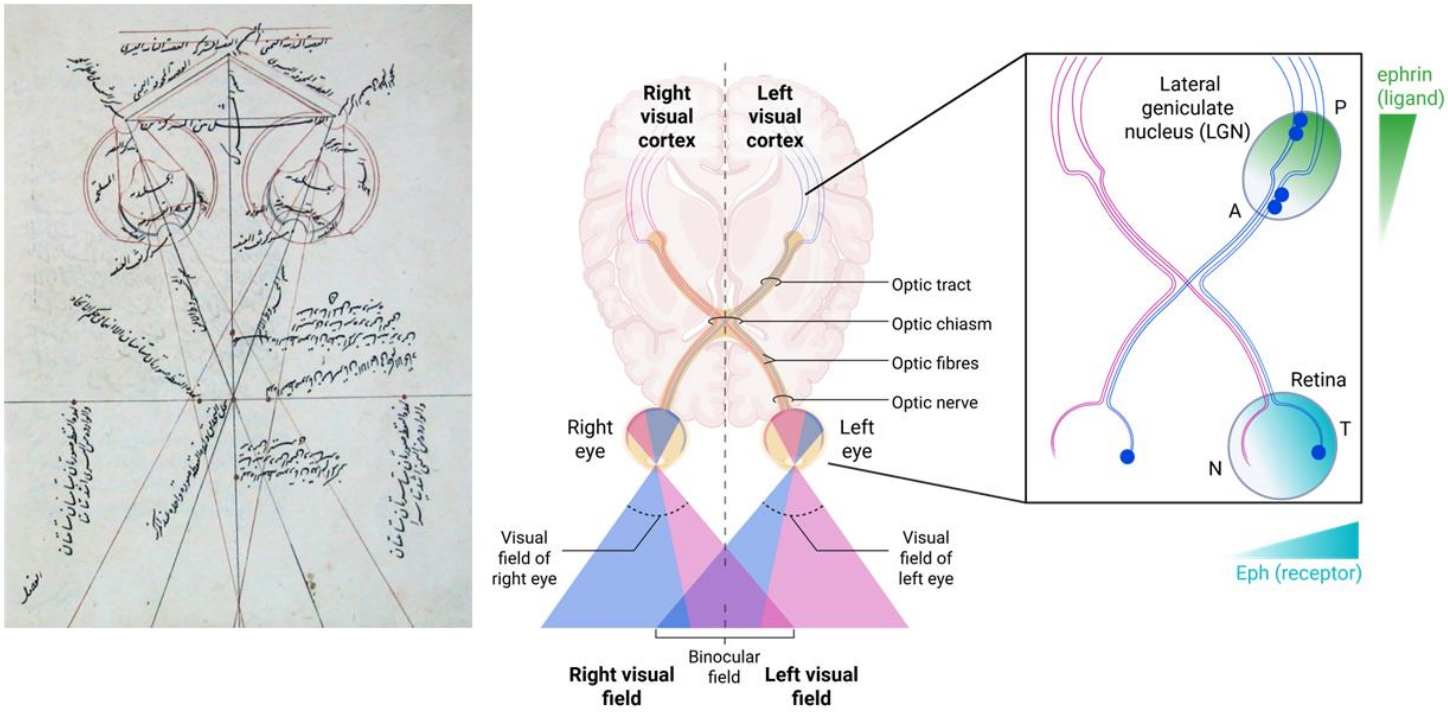

Figure 18.1 | (**left**) Formation of binocular vision through eye and optic chiasm as detailed by Kamal al-din Farsi in his Tankih al-Manazir (1309) building on Jurjani's anatomical observations and ibn al-Haytham's optics and vision theory. (**middle**) Binocular field forms at the intersection of vision coming from temporal retina of two eyes. (**right**) Retinal ganglion cells map to optic tectum in vertebrates (and LGN in mammals) as temporal retina nerves connect to anterior region and nasal nerves connect to the posterior region via ephrin ligand and Ephrin receptor gradients.

Another approach to the same problem: The compound eye

As noted in the introduction, many insects display quite a different solution for vision. Hexagonal packing of hundreds of ommatidia, individual vision units each sporting small crystalline lenses followed by photoreceptors and surrounded by pigment cells, forms the compound eye. Together, ommatidia function like an array of small eyes processing the image from different angles and sensitivities.

In Drosophila melanogaster, compound eye develops from eye imaginal disc in third instar larva. Each ommatidia are specified by eight neighbouring cells, with R8 in the middle. Those eight-celled groups in the eye disc epithelium are determined trailing behind the eye furrow where the cells are mechanically constricted and arrested for division. Hedgehog and Dpp (Shh and BMP homologs) signals trace the furrow to stop the division and prime the cells for neural fate. Repression from lateral Wingless (Wnt homolog) signal allows for R8 specification only in the middle of eye disc. Notch signalling based lateral inhibition ensures stable R8 commitment in non-neighbouring cells. As the furrow travels from posterior to anterior (towards antenna disc) successively R2 and R5 (laterally), R3 and R4 (to the anterior), R1 and R6 (to the posterior), and lastly R7 (most posterior) cells are fate specified around each R8 (in the middle). Surrounding each 8-cell group, cone cells (which deposit the lens) are also specified via gradual EGF signal expanding from R8. As the cell division returns with furrow moving further, those fate-committed rosettes give rise to individual ommatidia. To sum, signal gradients and dynamics in cartesian and radial coordinates, lateral inhibition, relay-type fate induction, and gradual return of cell division altogether pattern equally distanced and structured ommatidia precursors in the eye disc, which develop into insect compound eye.

CASE STUDY: Primitive Eyes?

What made Darwin's proposal of "evolution by natural selection" original was the idea that variation appearing in populations could be selected by environmental conditions, and that these selected variations could become fixed over many generations, gradually producing particular features of organisms. The eye was especially striking to Darwin because all its

parts seem to come together as a unit: the lens to collect light, the detection unit at the back called retina, and the other associated parts. These parts seem difficult to imagine arising independently. If there is no lens, the retina is of limited use; if there is no detector, the lens is of limited use. This is therefore one of the classic fundamental questions in biology, not only in developmental biology but also in evolutionary biology.

To grasp this apparent absurdity, it will be a good practice to visit different organismal solutions for light sensation we come across with in nature. Functionally, to detect light effectively, three things are needed. First, a lens or focusing structure is needed to collect light. Second, a detector is needed, positioned where the light is focused. Third, a dark backing or screening structure is needed behind the detector to absorb scattered light and preserve spatial information. Those three physical constraints on the evolving system can be readily observed in very simple organisms.

A fundamental light-sensing unit we should begin the discussion with are light-harvesting cyanobacteria. Those bacteria can focus light beam with their whole cellular bodies and move towards light source (away from where the light focuses). Certain marine plankton like dinoflagellates, which are single-celled organisms, on the other hand has specialized light-sensing organelles named "ocelloid"s. Ocelloids can be called "eye-like" in the sense that they comprise of components resembling complex eyes: a transparent spherical structure that acts like a lens, organized membrane layers enriched in light-sensitive molecules such as rhodopsins, and a dark pigmented region behind them that acts as a screen. Thus, even a single-celled organism can contain structures that perform an eye-like function. Ocelloids are likely put together through an assembly on endosymbiotic cyanobacteria.

A marine worm, polychaete, has a simplistic design larval eye, consisting of two to three cells. While a pigmented cell act as a shade, another photoreceptor cell with opsins can detect the light. Absorption by shading cell is essential to prevent any scattering photons to deceive the photoreceptor unit, that way the animal can sense the direction light is coming from. Despite the simplistic design, polychaete larvae successfully display phototaxis (in early stage to move up for feeding), light aversion (in late stage to reach the sea floor), and steering (by processing information from two eyes) behaviours using those eyes.

Nematodes like C. elegans lack a centralized nervous system and have no eyes. However, they display aversion to light, negative phototaxis, through light responsive neuronal cells located mainly in their heads. Notably, they have a taste family receptor protein, LITE-1, repurposed as a photosensor instead of Opsins often in animals with eyes. On the other hand, planarian flatworms have two eyes with optic cups, composed of pigment cells for shielding, and photoreceptor cells for light detection and display negative phototaxis by communicating those signals to their centralized brains through an optic chiasm much alike the vertebrate system. Furthermore, planaria have a body-wide sensory system which helps with reflexive light aversion.

A last example is Box jellyfish (Cubozoa), which possess complex lens eyes above their tentacles. These eyes contain all three components: lens, detector, and dark backing. They do not necessarily connect to a brain for image processing; instead, they may connect directly to muscles. Thus, visual information does not require a brain to be useful. If these three components are present together, light information can still influence behavior.

From this perspective, it becomes easier to think about how vision may have evolved. Photosensitive cells, especially photoreceptors associated with pigment cells, could behave as primitive eyes. In some species, only two cells—a pigment cell and a photoreceptor cell—can form a minimal visual organ. From such simple beginnings, more complex eye types could evolve in different lineages.

FURTHER READING

Historical perspective on the development and evolution of eyes and photoreceptors, W.J. Gehring, **The International Journal of Developmental Biology**, DOI: 10.1387/ijdb.041900wg.

Making senses: Development of vertebrate cranial placodes, G. Schlosser, **International Review of Cell and Molecular Biology**, 2010. DOI: 10.1016/S1937-6448(10)83004-7.

Eye Development and retinogenesis. W Heavner, L. Pevny, **Cold Spring Harbor Perspectives in Biology**, 2012. DOI: 10.1101/cshperspect.a008391.

Chapter 19: EVO-DEVO

"Any living cell carries with it the experiences of a billion years of experimentation by its ancestors. You cannot explain so wise an old bird in a few simple words."

– Max Delbrück, from "A physicist looks at biology", 1949.

"What characterizes the living world is both its diversity and its underlying unity."

– François Jacob, from "Evolution and tinkering", 1977.

In the last chapter of our inquisitive journey, we want to apply the fundamental design principles and key molecular mechanisms of developmental biology so far covered to the broader question of how continues to evolve on Earth. This chapter is not intended as a comprehensive review of evo-devo, but rather as a set of selected examples.

Evolutionary developmental biology asks how developmental mechanisms help explain the evolution of form. A historically influential idea was that "ontogeny recapitulates phylogeny", a phrase associated with Haeckel. Although his famed drawings and some other claims are heavily disputed, the idea here was based on a real observation: embryos of different vertebrate species can look remarkably similar at certain stages. Comparative embryologists in the nineteenth century noticed this. Karl von Baer even wrote in 1828: "I have two small embryos preserved in alcohol that I forgot to label. At present I am unable to determine the genus to which they belong. They may be lizards, small birds, or even mammals." This stage of strong similarity is called the phylotypic stage which has been covered in the book extensively.

In Haeckel's interpretation, development (ontogeny, formation of the individual) replayed evolutionary history (phylogeny, formation of the species). An organism was thought to pass through stages representing ancestral adult forms, which as we know today is a strong and incorrect claim. However, the observation of a conserved middle stage (in vertebrates) here is important. The hourglass model of vertebrate embryogenesis better captures the modern understanding. In the earliest stages, such as cleavage and blastula stages,

embryos of fish, frogs, birds, and mammals can look quite different. Then, in the middle of early embryogenesis, they pass through a relatively constrained stage in which they are remarkably similar. After that, they diverge again as development continues and species-specific features appear.

The phylotypic stage and hourglass model of development

At this phylotypic stage, several fundamental processes are happening. The embryo is forming segmented structures such as somites and hindbrain rhombomeres. Neurulation is occurring. All three major axes—dorsoventral, anteroposterior, and mediolateral—are established. The organizer has already produced the embryonic midline, including the notochord. The head-to-tail body plan is being coordinated with Hox gene expression. Altogether, this suggests that there is very strong selective pressure on the generalized vertebrate body plan. From an evolutionary perspective, there may be only a limited number of successful ways to establish a vertebrate organism.

We had discussed why the initial embryo shapes may vary in Chapter 5. But why is the phylotypic stage so constrained? Several reasons can be discussed. First, there are only a handful of signalling pathways used repeatedly in development. Second, many of these pathways involve diffusible proteins, and diffusion imposes spatial limits. A gradient cannot be arbitrarily long, *e.g.*, several millimetres, or arbitrarily short, less than couple 10 microns, and still function effectively. Thus, the physical scale of possible patterning in the tissues impose real developmental constraints. Third, cells are constrained by metabolic and degradation rates. Chemical reactions, protein synthesis, and protein turnover occur on limited time scales. Development cannot be made arbitrarily fast or slow without violating these biochemical constraints. Fourth, there are mechanical constraints at the molecular level. Cells rely on molecules such as catenin and cadherins to adhere and mechanically communicate (transmit forces and torques) to organize into tissues. These, too, impose constraints on how development can proceed. All these factors together may help force embryos through the phylotypic bottleneck of the hourglass model. In that sense, the vertebrate body plan is strongly constrained by both molecular and physical requirements.

Evolution of development

To discuss how development itself evolved, it is first necessary to discuss the evolution of multicellularity, since development in this sense applies to multicellular organisms. Let us focus on multicellular animals for simplicity. At the molecular level, one key point was that components such as collagen appear around 600 million years ago and helped cells stick together more effectively, enabling the formation of the first multicellular body plans. Around similar times, Hox genes also appear in the evolutionary record of proteins. These genes are central to multicellular body plans and likely already existed in ancestral bilaterian animals, perhaps as a set of around seven Hox genes.

Shortly after this, in evolutionary terms, came the Cambrian explosion, around 541 million years ago. This diversification may have unfolded over only about 10 million years, which is noticeably short compared to the history of life on Earth. During this interval, organisms acquired eyes, at least simple nervous systems or direct eye-to-muscle connections, and segmented body plans. Segmentation gave them modularity, which was advantageous for movement. Eyes and simple neural processing allowed organisms to detect and respond to one another more effectively, including chasing prey and escaping predators. All this likely contributed to the rapid diversification of animal forms. Majority of animal body plans (all of the extant) have been established during this period. Later evolution elaborated on these body plans.

The evolution of germ layers is another key step. Some organisms remained with only two germ layers, ectoderm and endoderm, whereas others evolved an additional mesoderm between them. Those that evolved mesoderm gained the capacity to form more complex bilaterally symmetrical body plans. Those without mesoderm generally remained more radially symmetrical.

An interesting example is the sea anemone *Nematostella vectensis*, a cnidarian. Although cnidarians are not bilaterians, *Nematostella* show bilaterally symmetrical features. This makes it informative for thinking about the ancestor shared by cnidarians and bilaterians. One interpretation is that bilateral tendencies may already have been present in the common ancestor and were elaborated differently in later lineages. Such evolutionary conclusions must be treated cautiously. In evolutionary biology, we cannot perform direct experiments on the

past. Instead, we here combine developmental biology, molecular evidence, comparative anatomy, and phylogenetic reasoning to propose hypotheses about what likely happened.

Mechanisms of novelty: Gene duplication and fusion

One molecular mechanism that help generate evolutionary novelty is gene duplication. Sometimes, during genome duplication events, an organism ends up with an extra copy of every gene. Because development is robust, such organisms can still develop. Once there are two copies of a gene, one copy may continue doing the original job while the other mostly freed of the original selective pressure can accumulate mutations more easily. These mutations may affect either the coding region or the regulatory regions such as promoters and enhancers.

1- Examples for the first is the relationship between β-catenin and γ-catenin. In vertebrates, γ-catenin, also called plakoglobin, appears to have arisen following a gene duplication event and later became repurposed to work in mechanical transduction rather than continuing in Wnt signalling as β-catenin does.

2- Example for the second is the even-skipped gene in fruit fly segmentation. The same even-skipped gene is expressed in different stripes, but each stripe is driven by different enhancers. Thus, the control regions evolved to allow different upstream transcription factors to activate the same gene in different positions.

A second important mechanism enabling evolutionary novelty is gene fusion, in which originally separate genes or domains become combined. Hedgehog was used as an example. Comparative sequence evidence suggests that a big non-functional domain of the hedgehog protein coding gene has likely existed in more ancestral organisms, as it persists in sponges and choanoflagellates. A later fusion or domain assembly event may have produced the complete hedgehog protein. Hedgehog later became a central developmental signal involved in specifying the midline and splitting structures into left and right as discussed through the book.

Cell–cell contact molecules

Let us now focus a bit deeper on cadherin and catenin molecules. Different animal clades possess different numbers of cadherins, protocadherins, cadherin-related proteins, and catenins, but the overall pattern suggests that these molecules co-evolved and became established at similar times, around 600 million years ago. These adhesion molecules were therefore among the molecular foundations that enabled multicellular body plans and later diversification in the Cambrian explosion. One interesting pattern concerns protocadherins. These are molecules that contain cadherin-like repeats but are distinct from the classical cadherins. They are relatively limited in many lineages, but in vertebrates they expand dramatically. This expansion may have been particularly important for the vertebrate body plan, including segmentation and nervous system complexity. For instance, a protocadherin, Papc, is especially important for separating segmented somites from one another.

The developmental signalling pathways

This brings us to the major developmental signalling pathways covered in the book: Wnt, Notch, Nodal, BMP, and Sonic hedgehog. From evolutionary perspective, Bilaterians possess all five. Cnidarians also possess all five. As we expand to porifera (*e.g.*, sponges) and placozoa, we lose Nodal and Shh out of the equation. Sponges are particularly interesting in this respect. The case for Hedgehog was discussed above. Although sponges are simpler-looking animals not showing the same degree of differentiated tissues as bilaterians or cnidarians, they still possess Notch, Wnt, BMP-related signalling, some homeobox genes, and some cell types that may be sensory-like. This suggests that the canonical developmental signalling toolkit evolved quite early in animal history, probably soon before the Cambrian explosion in between 550 and 600 million years ago. This developmental signalling toolkit was later reused repeatedly in more complex animals.

Common ancestors

The hypothetical common ancestor of the "true animals" (Eumetazoa) likely already had flagellated sperm, the ability to move, gastrulation, multiple germ layers—probably two—, an epithelium resting on a

basement membrane, a lined gut, and some kind of neuronal, muscular, and sensory systems. It also likely had anterior–posterior, dorsoventral, and mediolateral axes.

A second hypothetical ancestor is the ancestor of the bilaterians. Rather than being similar to protostomes (first mouth opening, *e.g.*, fruit flies), *Urbilateria* likely had its anterior–posterior and dorsoventral axes in a form similar to chordates (animals with notochord, *e.g.*, vertebrates) which are deuterostomes ("mouth second"). This arrangement would have a dorsal nervous system and ventral digestive system, and the protostome arrangement likely arose through dorsoventral axis inversion. This relates to earlier discussion of how BMP and related signalling systems are deployed in opposite orientations in vertebrates and arthropods.

Segmental metamerism

Segmented body plans have been central in evolution of diverse animal forms. Fruit flies, and other long germband insects, segment their body axis through the maternal-gene, gap-gene, pair-rule, and segment polarity cascade, and they do so almost simultaneously at blastoderm stage. However, this may be the exception rather than the rule. In many species, including vertebrates and short germband insects such as flour beetles, the body axis segments sequentially, one segment after another, from head to tail.

Comparative work suggests that the common underlying architecture of sequential segmentation consists of two things. First, there is a graded signal in the posterior, such as caudal in insects or FGF/Wnt in vertebrates. Second, there is a clock, often involving oscillatory expression of genes such as Notch pathway components. If a graded posterior signal retreats as the axis extends, and a clock interacts with that retreating zone, then new segments can be produced sequentially.

This logic appears broadly conserved. It is seen in vertebrates, in flour beetles, and even in plant root patterning, where a graded auxin signal interacts with a root clock oscillator. Thus, to generate sequential patterning, one generally needs a graded positional signal and a clock, and these two must interact.

What does this imply for the ancestral bilaterian *Urbilateria*? To address that, one must examine more basal species. Amphioxus already has segmented structures called somites, although it is not a vertebrate. Tunicates, by contrast, do not segment their body axis even though they possess graded posterior signals.

This suggests that the mere existence of a gradient is not enough; the gradient must retreat over the cells (and interact with a clock-like mechanism). Amphioxus appears to possess both a posterior signalling domain and something like an oscillatory clock gene. Thus, it is plausible that the common ancestor of bilaterians already had a graded posterior signalling system and perhaps also a clock, allowing some form of sequential segmentation.

Case Study: Remarkable Number of Vertebrae in Snakes

In vertebrates, the presomitic mesoderm elongates and somites form one by one as the segmentation clock ticks and as posterior FGF or Wnt signalling declines. The interesting question is how this system can be modified so much across species. Snakes can produce more than 300 vertebrae, whereas humans produce around 40 and zebrafish around 30. How can a robust segmentation system within each species produce such different outcomes in between them?

Gómez et al. (2008) addressed this question with comparative embryological work. Snakes have an unusually long presomitic mesoderm relative to the size of their somites. In other words, they maintain a long, unsegmented region near the posterior, allowing the segmentation clock to keep packing new somites over many cycles. Expression of clock targets in snakes shows many more stripes than in chick embryos, consistent with this prolonged segment-producing state.

One might think this is because the snake segmentation clock is much faster, but the data suggest that this is not the main explanation. Segmentation clock periods in corn snake,

chicken, and mouse are broadly comparable. What differs is the cell cycle. Snakes, and similarly long-tailed lizards such as whiptail lizards, have very long cell cycle durations. Thus, while the segmentation clock continues ticking at a fairly ordinary rate, cell generations proceed more slowly. Over the span of a given number of cell generations, more somites can therefore be produced. We shall note that many species seem to produce somites over roughly 10–20 cell generations. If those generations are slow, as in snakes, many more somites can be made. By contrast, zebrafish have only few cell generations to make somites but a faster segmentation clock, allowing them still to produce around thirty somites.

Amniotes may also benefit from the extra developmental support of the amniotic sac, allowing segmentation to continue for longer. However, this is not the end of the story. The species with the highest vertebral count are not from amniotes, not a snake, but a deep-sea eel, with around 750 vertebrae. Again, the explanation may involve an extremely slow cell cycle. Cultured eel cells can have cell cycle times on the order of 170 hours, which is extraordinarily slow. If the segmentation clock remains active while the cell cycle is slowed dramatically, then many more segments can form.

This provides a final evo-devo lesson: large changes in body form can arise by modifying the timing of conserved developmental systems rather than inventing entirely new ones. Changing the relationship between the cell cycle and the segmentation clock can greatly increase vertebral number while leaving the underlying segmentation mechanism largely intact.

FURTHER READING

Evolution and tinkering. F. Jacob, **Science**, 1977. DOI: 10.1126/science.860134.

The Plausibility of Life: Resolving Darwin's Dilemma. J.C. Gerhart, M. Kirschner, **Yale University Press**, 2006. ISBN: 0300119771.

The evolution of developmental mechanisms. **Development (supplement issue)**, 1994. https://journals.biologists.com/dev/issue/1994/Supplement.

Control of segment number in vertebrate embryos. C. Gomez *et al.*, Nature, 2008. DOI: 10.1038/nature07020.

A design logic for sequential segmentation across organisms. M.F. Simsek, E.M. Özbudak, **The FEBS Journal**, 2023. DOI: 10.1111/febs.16899.

Case Study Index

Figures Index

Concepts Index