Spatial dynamics of intestinal homeostasis

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. H. Bijl,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
donderdag 7 mei 2026 om 10:00 uur

door

Max Alphons BETJES

Dit proefschrift is goedgekeurd door de (co)promotoren.

Samenstelling promotiecommissie:

Rector magnificus,	voorzitter
Prof. dr. ir. S. J. Tans,	Technische Universiteit Delft/AMOLF, promotor
Prof. dr. J.S. van Zon,	Universiteit Leiden/AMOLF, copromotor

Onafhankelijke leden:

Prof. dr. B. Rieger	Technische Universiteit Delft
Prof. dr. C.P. Broedersz	Vrije Universiteit
Dr. D. Delacour	Aix Marseille Université, Frankrijk
Dr. N.A. Kurniawan	Technische Universiteit Eindhoven
Dr. R. Levayer	Institut Pasteur, Frankrijk
Prof. dr. C. Joo	Technische Universiteit Delft, *reservelid*

Het onderzoek beschreven in dit proefschrift werd uitgevoerd te AMOLF, Amsterdam en gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)

Kernwoorden: Darmepitheel, organoïden, live microscopie, celtracking

Druk: Ridderprint

Omslag: De korenoogst, Pieter Bruegel de Oude, Metropolitan Museum of Art, New York

Een elektronische versie van dit proefschrift is verkrijgbaar op: http://repository.tudelft.nl

Contents

Summary **ix**

Samenvatting **xiii**

1 Introduction **1**

2 Cell Tracking for Organoids: Lessons from Developmental Biology **9**

2.1 Introduction 10

2.2 Automated cell tracking 11

2.3 Cell tracking combined with fluorescent markers 13

2.4 Cell tracking combined with end-point measurements 15

2.5 Discussion 16

3 Cell tracking with accurate error prediction **19**

3.1 Introduction 20

3.2 Results 23

3.3 Discussion 36

3.4 Methods 38

3.5 Supplementary figures 58

4 Organoid cell fate dynamics in space and time **93**

4.1 Introduction 94

4.2 Results 95

4.3 Discussion 104
4.4 Materials and Methods 107
4.5 Supplementary figures 113

5 Sequential expression of tuft cell subtypes along the crypt-villus axis 125
5.1 Introduction 126
5.2 Results 127
5.3 Discussion 144
5.4 Methods 147
5.5 Supplementary figures 154

6 Mechanically driven checkerboard-patterning in the intestinal crypt 167
6.1 Introduction 168
6.2 Results 169
6.3 Conclusions 183
6.4 Methods 185
6.5 Supplementary Figures 189
6.6 Appendix: Model description 193

7 Discussion 205

Acknowledgements 221

About the author 223

Publications 225

Bibliography 227

Summary

The intestinal epithelium is constantly renewing itself, generating all its different cell types from a dedicated pool of stem cells. These stem cells are located at the bottom of so-called crypts, invaginations lying in between the villi, the finger-like protrusions into the intestinal cavity. The villi are where cells after migrating out of the crypts will die. Cells within intestinal epithelium thus have to coordinate changes in cell type with changes in position, with respect to the crypt-villus axis as well as with respect to their direct neighbors to maintain both global and finer scale patterns. How cells achieve this coordination is not yet understood. In this thesis, live tracking of cells in organoids (3D structures grown from stem cells) is used to investigate the relationship between cell fate and cell movement.

The introduction, **chapter 1**, provides an overview of the different cell types in the intestinal epithelium, their positions and what is known about their differentiation process. Major open questions regarding the interaction between cell fate and movement are discussed as well as the current bottlenecks in answering these questions.

In **chapter 2,** we discuss the benefits of our cell tracking approach, outlining its advantages over other methods. In particular, we compare it to genetic lineage tracing using inducible tracer proteins, where lineage structure can only be identified indirectly from analysis of the final state. In contrast cell tracking allows researchers to follow cells throughout time and space at the single cell level. We show how cell tracking has been combined in different developmental systems with other molecular biology techniques, such as antibody staining and fluorescent labelling, to further enrich the tracking data. This chapter thus provides an overview of the state-of-the-art at the start of this thesis work.

The size of the 3D imaging datasets needed for tracking intestinal organoid differentiation renders manual analysis impossibly time consuming. Therefore, machine learning powered semi-automated tracking methods have been developed. In **chapter 3** we present our contribution to the cell tracking field by tackling a major obstacle in its broader adoption, namely its lack of interpretability. Current cell tracking algorithms function largely as black

boxes, producing cell tracks without a reliable indication of tracking quality. This lack necessitates extensive manual review before downstream analysis can be performed. We present a conceptually new tracking framework where track predictions are accompanied by error rate estimates. This is done by integrating calibrated neural network predictions on individual cells and timepoints in a statistical framework that combines them into error rates without loss of calibration. These error rates can guide manual review or allow for automatic filtering out of potential errors. The latter, for the first time, facilitates fully automated analysis of key lineage dynamics parameters based on the remaining high-confidence data.

In **chapter 3** we combine cell tracking with immunostaining to identify all major intestinal cell types at the experimental endpoint. This combines spatially resolved lineage information (cell tracking) with cell type information (staining). In our analysis we find that cells generally commit to their fate early and still within the stem cell zone. Only afterwards do the cells sort to their proper position on the crypt-villus axis. Cells also directly commit to their final state without passing to progenitor states shared with other types. We term this the 'commit-then-sort' model of cell differentiation.

In **chapter 4** we combine cell tracking with single cell sequencing to study one particular cell type; the tuft cell. This is a very rare cell type, whose differentiation and function are still poorly understood. Using single cell sequencing we identify various subtypes of tuft cells that seem to lie on a continuum with respect to the crypt villus axis. Using cell tracking of fluorescent markers of the various subtypes in organoids, we show that subtypes transition into each other while aging and moving from the crypt to the villus. By modifying spatial cues (mainly in the form of BMP-signaling) we can then control the tuft subtype distribution, allowing for more detailed functional characterization of these subtypes.

In **chapter 5**, we look at the role of mechanics in the patterning of intestinal cell types. We specifically focus on the crypt, where secretory cells are surrounded by non-secretory cells, most prominently at the crypt bottom where secretory Paneth cells form a so-called 'checkerboard' pattern with the stem cells. These patterns are typically thought to be caused by lateral inhibition through Notch signaling, in which a secretory cell inhibits that fate in its neighbors. How can such patterns be maintained in dynamical systems like the intestinal epithelium where cells are constantly dividing? Using modelling and cell tracking, we show that the rate of cellular rearrangement is too fast for lateral inhibition to maintain the pattern. Instead, our tracking suggests that the pattern is maintained by the fast separation of secretory neighbors. Indeed, when we use laser-ablation to create secretory neighbor pairs, the pattern is re-established through secretory neighbor separation without changes in cell fate. This process seems driven by differences in actomyosin contractility between the stem-stem and stem-Paneth surfaces. Our conclusions are a major departure from previous models that assumed cell signaling, rather than cell mechanics, as the driver of this spatial patterning.

In the final chapter, we summarize and comment on the major findings of this thesis. Various extensions for our cell tracking methodology are offered alongside some speculation of where

the field is or should be headed. An interesting application of the methods developed in this thesis would be to look at inflammation, where cell differentiation seems to differ markedly. I suggest that studying the differences between healthy and inflamed differentiation pathways might also shed light on the constraints under which the system operates, what features remain the same? The same is true for comparisons with different organs like the skin, lung and mammary gland epithelium that I suggest in this chapter.

Samenvatting

Het het epitheel van de darm vernieuwt zichzelf voortdurend, waarbij al de verschillende celtypen door een gespecialiseerde pool van stamcellen wordt gegenereerd. Deze stamcellen bevinden zich aan de bodem van zogegenaamde crypten, instulpingen tussen de lange vingervormige uitsteeksels die de darmwand bekleden. Zo een 'vinger' wordt op zijn beurt een villus genoemd en het is hier dat de oude epitheel cellen afsterven, nadat ze vanuit de crypten daar naartoe zijn gemigreerd. Tijdens die reis differentieren de cellen van stamcel tot specifieke celtypen. Cellen in het darmepitheel moeten dus veranderingen in celtype coördineren met veranderingen in positie, zowel ten opzichte van de crypt-villus-as als ten opzichte van hun directe buren, om zowel globale als lokale patronen in stand te houden. Hoe cellen precies deze coördinatie tot stand brengen is nog onbekend. In dit proefschrift wordt live tracking van cellen in organoïden (3D-structuren gekweekt uit stamcellen) gebruikt om de relatie tussen celtype en celmigratie te onderzoeken.

De inleiding, **hoofdstuk 1**, geeft een overzicht van de verschillende celtypen in het darmepitheel, hun posities en wat bekend is over hun differentiatieproces. Belangrijke open vragen over de interactie tussen het uiteindelijk celtype en het traject dat cellen afleggen worden besproken, evenals de huidige knelpunten bij het beantwoorden van deze vragen.

In **hoofdstuk 2** bespreken we de voordelen van onze celtracking-aanpak ten opzichte van andere methoden. In het bijzonder vergelijken we onze techniek met genetische tracering van cellen door middel van induceerbare tracer-eiwitten, waarbij de stambomen en migratietrajecten van cellen alleen indirect kunnen worden afgeleid door de locaties van de gelabelde cellen te analyseren op het experminetel eindpunt. Celtracking daarentegen maakt het mogelijk om cellen in de tijd en ruimte te volgen op het niveau van individuele cellen. We laten zien hoe celtracking in verschillende ontwikkelingsbiologische modelsystemen is gecombineerd met andere moleculair-biologische technieken, zoals immounohistologie en fluorescente labeling van eiwitten, om de trackingdata verder te verrijken. Dit hoofdstuk biedt daarmee een overzicht van de 'state-of-the-art' aan het begin van dit proefschrift.

De omvang van de microscopie datasets die nodig zijn om de differentiatie van darm-organoïden te volgen, maakt handmatige analyse onuitvoerbaar tijdrovend. Daarom zijn in de afgelopen jaren verschillene door artificiele intelligentie aangedreven semi-geautomatiseerde trackingmethoden ontwikkeld. In **hoofdstuk 3** presenteren we onze bijdrage aan het veld van celtracking waarin we een belangrijke belemmering voor de bredere toepassing van deze methodes aanpakken, namelijk het gebrek aan interpreteerbaarheid. Huidige celtrackingalgoritmen functioneren grotendeels als 'black boxes' en produceren celtracks zonder betrouwbare indicatie van de trackingkwaliteit. Dit maakt arbeidsintensieve handmatige controle noodzakelijk voordat verdere analyse kan plaatsvinden. Wij presenteren een conceptueel nieuwe trackingraanpak waarin voorspellingen van celtracks worden aangevuld met schattingen van de foutkansen voor elke stap in de track. We doen dit door gecalibreerde neurale-netwerk voorspellingen op het niveau van individuele cellen en tijdspunten te integreren in een statistisch raamwerk dat deze voorspellingen combineert tot foutkansen, zonder dat de callibratie verloren gaat. Deze voorspelde foutkansen kunnen vervolgens handmatige controle sturen of het automatisch uitfilteren van potentiele fouten mogelijk maken. Dit laatste maakt voor het eerst volledig geautomatiseerde analyse van belangrijke parameters van celdifferentiatie mogelijk.

In **hoofdstuk 3** combineren we celtracking met immunohistologie om alle belangrijke celtypen in de darm aan het einde van het cell tracking experminent te identificeren. Deze methode combineert daarmee ruimtelijke en temporale informatie (celtracking) met informatie over celtypen (histologie). In onze analyse constateren we dat cellen zich over het algemeen vroeg committeren aan een specifiek celtype, nog binnen het stamcelcompartiment. Pas daarna sorteren de cellen naar hun juiste positie op de crypt-villus-as. Wij noemen dit het 'commit-then-sort'-model van celdifferentiatie.

In **hoofdstuk** 4 combineren we celtracking met 'single-cell sequencing' om één specifiek celtype te bestuderen: de tuftcel. Dit is een zeer zeldzaam celtype waarvan differentiatie en functie nog slecht begrepen zijn. Met behulp van 'single-cell sequencing' identificeren we verschillende subtypen van tuftcellen die op een continuüm lijken te liggen ten opzichte van de crypt-villus-as. Door celtracking van fluorescente markers voor de verschillende subtypen in organoïden laten we zien dat deze subtypen in elkaar overgaan naarmate tuft cellen ouder worden en zich van de crypt naar de villus verplaatsen. Door ruimtelijke signalen (voornamelijk in de vorm van BMP-signalering) te modificeren, kunnen we vervolgens de verdeling van tuft-subtypen sturen, wat een gedetailleerdere functionele karakterisering van deze subtypen mogelijk maakt.

In **hoofdstuk 5** onderzoeken we de rol van mechanica in de onderlinge ordening van intestinale celtypen. We richten ons specifiek op de crypt, waar secretoire cellen stevast worden omgeven door niet-secretoire cellen, met name onderaan de crypt waar secretoire Panethcellen samen met stamcellen een zogenoemd 'dambordpatroon' vormen. De vorming van deze patronen worden doorgaans toegeschreven aan laterale inhibitie via Notch-

signalering, waarbij een secretoire cel het secretoire differentiatieproces in zijn buren onderdrukt. De vraag blijft dan, hoe dergelijke patronen in stand worden gehouden in dynamische systemen zoals het darmepitheel, waarin cellen voortdurend delen en bewegen? Met behulp van een theoretisch model en celtracking laten we zien dat cellen in het darmepitheel te mobiel zijn voor laterale inhibitie om het patroon in stand te houden. In plaats daarvan suggereert onze tracking dat het patroon wordt behouden doordat secretoire buurparen snel uit elkaar bewegen. Wanneer we met laserablatie artificieel secretoire buurparen creëren, wordt het patroon inderdaad hersteld door het uit elkaar bewegen van secretoire buren zonder dat er (de)differentiatie optreedt. Dit proces lijkt te worden aangedreven door verschillen in actomyosinecontractiliteit tussen stamcel-stamcel en stamcel-Paneth-oppervlakken. Onze conclusies wijken sterk af van eerdere modellen die veronderstelden dat celsignalering, en niet celmechanica, de drijvende kracht was achter deze patroonvorming.

In het laatste hoofdstuk vatten we de belangrijkste bevindingen van dit proefschrift samen. Verschillende uitbreidingen van onze celtrackingmethodologie worden voorgesteld, gevolgd door enige speculatie over waar het celtracking veld zich nu bevindt en hoe het zich zou moeten ontwikkelen. Een interessante toepassing van de in dit proefschrift ontwikkelde methoden is het bestuderen van ontsteking in het darmepitheel, waarbij celdifferentiatie duidelijk anders lijkt te verlopen. De verschillen tussen gezonde en ontstoken celdifferentiatie kunnen in potentie inzicht geven in de randvoorwaarden waarin het systeem moet voldoen: welke kenmerken blijven altijd gelijk? Hetzelfde geldt voor vergelijkingen van het darmepitheel met andere organen, zoals de huid-, long- en melkklier-epithelia, die ik in dit hoofdstuk voorstel.

1 Introduction

Developmental biology is a question of cells of the right type, being in the right place, at the right time. It is thus natural to ask how cells coordinate changes in their identity with changes in their spatial position. As extremes, we could imagine the cells being highly autonomous agents, which only after commitment to a certain cell fate migrate to their proper place in the organism. Equally possible is a scenario of cells constantly responding to environmental cues and adapting their fate to their spatial position.

These questions regarding cell autonomy versus external control and determinism versus plasticity are among the oldest in developmental biology. They have been studied extensively in embryos, since the early twentieth century, when scientists like Spemann and Mangold would cut and paste pieces of newt embryo into new locations[1]. Would these transplanted cells change their behavior to fit their new context or continue on their own path?

But the interplay between cell identity and position is interesting beyond development as well. In adult organisms, cells constantly die from damage or old age and have to be replaced, while keeping the tissue functional and intact. Particularly fascinating cases of this homeostasis are rapidly dividing tissues like the skin, blood cells in the bone marrow, and, fastest among them, the intestine. The human small intestinal lining is regenerated every 5 days, generating a daily 200 grams of dead cells in the process[2]. This enormous task is further complicated by the fact that the intestinal epithelium consists of many different cell types, with completely distinct functions. These are thus systems where both the cells' positions and functions must rapidly change to facilitate self-renewal. How are organs able to maintain order while being continuously turned over?

Indeed, maintaining homeostasis is no trivial task and many intestinal diseases can be thought of as defects in homeostasis. Chronic inflammatory diseases, like Crohn's disease, are often accompanied by marked changes in how cells differentiate, leading to both changes in cell type composition and functionality[3,4]. Similarly, cancer, in its early stages, is to an extent just

cells escaping normal homeostatic control mechanisms. It is then no surprise that the altered homeostasis caused by chronic inflammation is a significant risk factor in tumor formation[5,6].

The problem of how to achieve robust homeostasis is also interesting because an optimal strategy is not readily clear. When returning to the two extremes introduced earlier, contrasting cell autonomy with external control, both seem to carry benefits. Allowing cell plasticity influenced by outside cues might make the system more robust while early cell autonomous decision making seems simpler and more efficient. Both these scenarios also lead to obvious downstream questions. If the cells have to adjust, how do they accurately sense their position, and if a cell's fate is fixed how does it ensure it moves to the right place? Optimal solutions seem to be highly context dependent, as can already be observed when looking at the most basic of cell fate decisions; when to stop dividing. During intestinal development, cells initially have similar proliferative potential with fate decisions made only later[7], while during adult homeostasis proliferative potential differs between different compartments with fate decisions made early[8]. Both strategies can be thought of as optimal depending on how we define their goals and constraints.

In this way, studying intestinal fate commitment and its interplay with cell movement may translate into a better understanding of what is evolutionary important and under what biological constraints the system operates. In that context, the mouse intestine can serve as a model for homeostasis within tissues, much like the Spemann and Mangold's newts did for embryonic development.

Lineage dynamics in the small intestine

How does one go about studying the mutual dependencies between environment and cell fate during homeostasis? Cutting and pasting cells around, while difficult in embryos, is generally impossible in adult animals. Luckily then, molecular biology has more recently provided new tools to study fate commitment. Lineage tracing is one such technique, which has proven key in resolving intestinal differentiation processes. Here, an organism is genetically modified such that the expression of a certain gene within a cell induces the permanent expression of a tracer marker (e.g. a fluorescent protein) in all its progeny. When sacrificing the animal, it can then be seen if a cell arose from a cell expressing the gene of interest, which could for instance be a marker for a putative stem cell population. By making the tracer induction dependent on an external signal as well (e.g. a drug), the expression of the tracer marker can be further restricted to cells whose ancestor expressed the protein of interest at an experimentally controlled timepoint[9]. Through the engineering of mouse lines for many proteins that are expressed by different cell types and in different compartments in the intestinal epithelium, scientists have mapped the developmental potential of these cells; what kind of progeny are they able to make and which places in the organ does their progeny reach.

The results of these and other experiments paints the following picture[10]; the small intestinal epithelium has finger-like protrusions called villi which house fully differentiated non-

dividing cells. These can be absorptive (enterocytes), tasked with nutrient up-take, or secretory, releasing mucus (goblet cell), hormones (enteroendocrine cells) or cytokines (tuft cells). Within these cell types there can be subtypes, for example, in the enteroendocrine lineage, where different subtypes are responsible for producing different hormones. In between the villi are invaginations called crypts, that are around 20 cells deep. The bottom of these crypts houses a small population of rapidly dividing stem cells (~30 cells). They are interspersed with another secretory type, the Paneth cells, which release antimicrobial compounds and stem cell maintaining signals. Above them in the crypt are the transit-amplifying cells, containing fast dividing absorptive progenitors (transit-amplifying cells) this time interspersed with secretory progenitor cells. Lineages thus start at the crypt bottom with a stem cell and from there its progeny moves up through the crypt and into the villus, where cells die and are shed from the epithelium. During this migration the cells differentiate from stem cells to mature absorptive or secretory types.

Organoids

Intestinal differentiation is a strongly hierarchical process, with the stem cell both generating all other cell types and regenerating themselves. Notably, intestinal stem cells are even able to do this *ex vivo*, when cultured in gels and media that mimic the in vivo environment well enough[11]. In this case, stem cells can generate three dimensional structures called organoids. These consist of a single cell layered epithelium containing all major cell types surrounding the lumen, which corresponds to the inside of the intestinal tube *in vivo*. Intestinal organoids, initially derived from biopsies, can be propagated indefinitely by fragmenting them and reseeding them in fresh gels. Crucially, the organoids not only recapitulate the intestinal cell types but also its spatial organization: crypts house the stem cells, Paneth cells and progenitor cells, while branching off from a central villus region containing enterocytes and goblet cells[10]. The way cell type and position are attuned to one another in the epithelium thus allows for a remarkable degree of self-organization in a way that is still only partially understood.

Open questions

Lineage tracing experiments have thus provided us with a global picture of intestinal homeostasis. We know the location of the major cell types along crypt-villus axis and have a sense of where the cell type transitions take place. Still, how the differentiation process is controlled often remains unclear, especially regarding the causal relation between cell position and cell fate.

For instance, it remains unknown where and when a stem cell's decision to differentiate is made, and most pertinently, whether this occurs; before or after it has left the stem cell zone. It is therefore unclear if the decision to differentiate precedes the exit of this zone or is rather the result of this exit. The same is true for the location and timing of other fate decisions,

such as the choice between a secretory and absorptive fate. In the absence of such information, it is hard to say if cell position controls cell fate or the other way around.

Another open question pertains to the exact nature of these fate decisions. These are often conceptualized as a series of binary choices, the first of which is a decision between the absorptive and secretory fate. The secretory progenitors would then make further decisions towards one of the major secretory cell types. Alternatively, one could imagine stem cells choosing a single fate directly, without going through intermediate states. The existence of a secretory progenitor is attractive not only on the basis of their common 'secretory' functionality, but also because of their shared dependence on the Notch/Delta signaling pathway[10,12]. Still, on the basis of lineage tracing and genetic perturbation experiments people have reached opposing conclusions about the nature of such a progenitor, including whether it even exists at all[9,13].

An as of yet less explored area is the relation between space and subtype, for cell types such as enteroendocrine or tuft cells that occur as distinct subtypes. RNA sequencing experiments have shown region specific enrichment of different subtypes of entero-endocrine and tuft cells along the crypt-villus axis[14,15]. The mechanisms by which this zonation is regulated remain largely unclear, especially in the case of the tuft subtypes.

Lastly, little is known about how cells regulate their relative positions within a certain compartment. For instance, secretory types are spatially isolated from each other, separated by absorptive cells or stem cells. This pattern is thought to emerge through a process called lateral inhibition, whereby cells committed to the secretory fate instruct their neighbors to acquire an opposite fate [16]. However, it is not obvious that such a mechanism, which relies on signaling between neighbors, would work highly for the highly proliferative and migratory cells in the intestinal epithelium, where a cell's neighbors might constantly change.

Current bottlenecks

The question posed above are hard to answer with current methods as they require lineage trees at single cell resolution in combination with spatial information to precisely pinpoint the locations of fate decisions. In lineage tracing experiments this level of detail is difficult to achieve because of its low temporal resolution. Such experiments fundamentally provide only two timepoints: the moment the tracer protein expression is induced and the moment the animal is sacrificed to analyze the clones. In between these timepoints cell migration and lineage structure can only be inferred statistically. Furthermore, lineage tracing presupposes significant prior knowledge about the system since one needs known markers for a specific subset of cells to induce tracer expression[9]. It is therefore poorly suited for more exploratory questions.

More recently single-cell sequencing has opened up a new avenue of studying differentiation. Instead of relying on a handful of known molecular markers, cell types can now be identified by clustering on the basis of their expression of ~20.000 different mRNAs. In the intestine,

single-cell sequencing has already revealed a previously unappreciated degree of variability within mature cell types[17], but at high enough throughput it should also be able to identify short-lived early progenitor types. With this comes the promise of ordering all the cells in the dataset from stem cell to the mature differentiated types on a lineage tree, such that one can simply read off the developmental pathways[18]. Unfortunately, reconstructing a dynamic picture from what is essentially a sequence of snapshots in time is not trivial, even when huge datasets are available[19,20]. More fundamentally, during sequencing all spatial information is lost, so even if the developmental trajectories could be perfectly reconstructed, their dependence on spatial factors would remain tenuous. With such techniques, the ways in which a cell's fate can influence its position, for instance through cell migration or sorting, are completely hidden from us.

Cell tracking in organoids

In this thesis, we use a method that does provide high-resolution data in space and time: cell tracking in organoids. This technique relies on two major scientific developments that happened in the past decade. The first are organoids themselves, whose *ex-vivo* nature makes them ideally suited for high-resolution live imaging, and following cells through time while they differentiate. The second development concerns the use of machine learning in image analysis and cell tracking in particular, that can greatly alleviate the burden associated with manual analysis. A significant part of this thesis is dedicated to the improvement of such cell tracking methods. It is this combination of 3D imaging, organoid culture and automated analysis, that enables us to for the first time study the spatiotemporal dynamics of intestinal differentiation and patterning at the single cell level.

Thesis overview

In **chapter 2,** which is based on a review article, we discuss the benefits of the cell tracking approach in more detail, outlining its advantages over other methods and explaining how cell tracking can be combined with other molecular biology techniques, such as antibody staining and fluorescent labelling. We also review interesting work done with cell tracking in developmental systems and the lessons they hold for organoid research. It thereby provides an overview of the state-of-the-art at the start of this thesis work.

Chapter 3 concerns the main bottleneck of cell tracking experiments: the practical task of following hundreds of cells over multiple day long experiments. The size of these 3D imaging datasets renders manual analysis impossibly time consuming. Therefore, machine learning powered semi-automated tracking methods have been developed. Unfortunately, these algorithms function largely as black boxes, producing cell tracks without a reliable indication of tracking quality. This lack necessitates extensive manual review before downstream analysis can be performed. We present a conceptually new tracking framework where track

predictions are accompanied by error rate estimates. These can guide manual review or even, for the first time, facilitate fully automated analysis of key lineage dynamics parameters.

Next, we combine cell tracking with subsequent immunostaining for all major intestinal cell types **(chapter 4).** This combines spatially resolved lineage information (cell tracking) with cell type information (staining). In our analysis we find that cells generally commit to their fate early and only afterwards sort to their proper position on the crypt-villus axis.

In **chapter 5** we zoom in on one particular cell type: the tuft cell. Using single cell sequencing and cell tracking of fluorescently labeled tuft cells in organoids, we studied how their subtypes are regulated. Here, we reveal an important regulatory role for position along the crypt-villus, with changes in the corresponding spatial cues driving cells from one tuft subtype to another.

Finally, in **chapter 6**, we shift the focus from how fate commitment interacts with location along the crypt-villus axis, to the ordering of cell types within a spatial domain. We specifically focus on the crypt bottom, where Paneth cells are surrounded by stems cells in a checkerboard pattern. How can such patterns be maintained in dynamical systems where cells are constantly dividing? Using modelling and cell tracking, we show that the rate of cellular rearrangement is too fast for cell differentiation to maintain the pattern. Using targeted laser-ablation, we then show that the pattern is instead maintained mechanically, by the fast physical separation of already differentiated neighboring Paneth cells. Our conclusions are a major departure from previous models that assumed cell signaling, rather than mechanics, as the driver of this spatial patterning.

2 Cell Tracking for Organoids: Lessons from Developmental Biology

Organoids have emerged as powerful model systems to study organ development and regeneration at the cellular level. Recently developed microscopy techniques that track individual cells through space and time hold great promise to elucidate the organizational principles of organs and organoids. Applied extensively in the past decade to embryo development and 2D cell cultures, cell tracking can reveal the cellular lineage trees, proliferation rates, and their spatial distributions, while fluorescent markers indicate differentiation events and other cellular processes. Here, we review a number of recent studies that exemplify the power of this approach, and illustrate its potential to organoid research. We will discuss promising future routes, and the key technical challenges that need to be overcome to apply cell tracking techniques to organoid biology.[1]

[1] The content of this chapter is published as: Betjes MA*, Zheng X*, Kok RNU*, van Zon JS, Tans SJ. Cell Tracking for Organoids: Lessons from Developmental Biology. *Front Cell Dev Biol.* 2021 (*Equal contribution)

2.1 Introduction

While the development and maintenance of organs is one of the most fundamental problems in biology, our understanding of it is far from complete. A hallmark of this process is the differentiation of cells in time, in terms of proliferative potential and cell type, with individual cells giving rise to complex lineages that organize in space to shape tissues and organs. Thus far, these differentiation dynamics have often been studied using the lineage tracing method[21-23]. Here, cells are labeled with a heritable marker such as fluorescent genes or a genetic barcode, for instance using Cre-Lox recombination[24,25] or lentiviral transduction[26-28]. This label can be detected in progeny after a certain period by fluorescence microscopy or single-cell sequencing and hence allows inference of genealogical relations between cells.

However, *lineage tracing* does not yield complete lineage trees nor provide information on the temporal dynamics of cells, such as their movements, growth rates, transient signaling, and timing of differentiation events (**figure 2.1**), which limits progress on many important questions. For instance, it remains largely unclear when and where cell fates are actually set, whether differentiation is either a consequence or a cause of spatial organization, how size and shape homeostasis is achieved, or how lineage dynamics are remodeled upon injury or disease. We also know little about the possible interplay with cellular metabolism, and the plethora of molecular signals from adjacent cells, for instance from the immune system. Elucidating the spatio-temporal dynamics is central to resolve these crucial issues and to elucidate the organizational principles of organ development[29].

A different technique, here referred to as *cell tracking* has the potential to reveal these developmental dynamics. As opposed to *lineage tracing* based on static snapshots, cells are here followed in real time over multiple generations, which can thus provide temporal dynamics, complete lineage trees, as well as spatial organization and cellular movements. Furthermore, *cell tracking* can readily be combined with the large spectrum of microscopy techniques that have been developed to study cell biology. For instance, the expression of fluorescent proteins directly identify differentiation events, cell-cycle progression, cytoskeletal structures, the dynamics of key molecular signals like Wnt or Notch, while FRET sensors may detect more rapidly changing signals such as calcium and metabolites[30-32].

Cell tracking has been applied extensively to study the early development of embryos, at increasing levels of sophistication[33-35], but poses challenges to the study of organs, given the challenges of time-lapse microscopy deep within tissues at later stages of development, even as intra-vital imaging is possible at lower resolution and throughput[36]. Yet, in recent years organoids have emerged as a model system for studying development and disease at the cellular level, including patient derived systems, which are ideally suited for this approach. Organoids are self-organizing cellular assemblies, which are grown *in vitro* and recapitulate organ structure and functionality to a striking degree[37,38]. Due to their *in vitro* nature, the growth and maintenance of organoid tissue can be observed directly by time-lapse

microscopy[39]. Yet, cell tracking approaches have only scarcely been applied to organoids thus far.

The aim of this review is to discuss the potential of cell tracking approaches for organoid science, as well as its technical challenges. We will do so by focusing on developmental systems ranging from 2D cell cultures to developing embryos, which have been studied extensively by cell tracking methods, and illustrate the scientific questions that it can uniquely address.

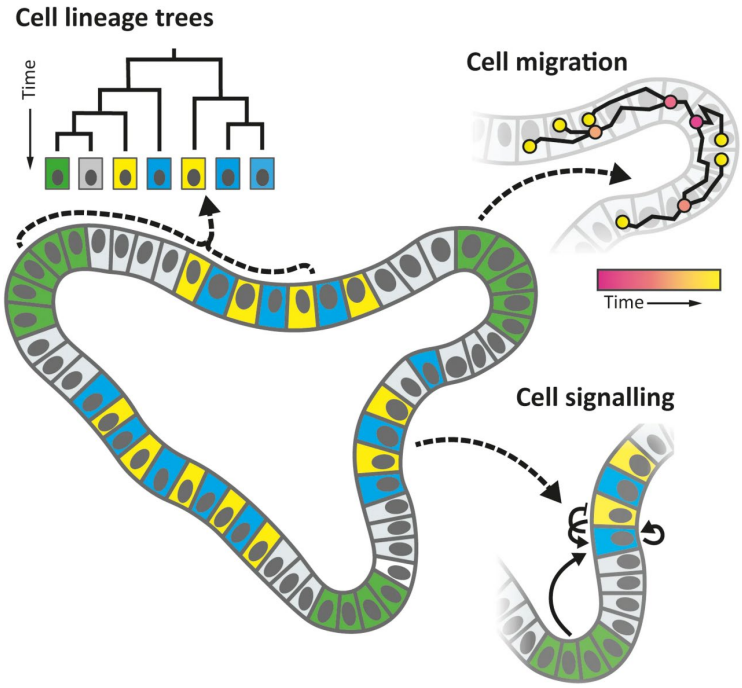


Figure 2.1: Organoid properties that can be studied by 'cell tracking', a technique in which (nearly) all cells are followed in time over multiple generations, using 3D time-lapse microscopy and automated image analysis, along with a host of fluorescence detection methods. Tracked cell positions allow one to reconstruct lineage trees, follow migration, growth and division dynamics, while fluorescence reporters may be used to study differentiation events, signaling pathways, and metabolic states, which are key to understand cellular decision-making during development.

2.2 Automated cell tracking

Automated tracking of individual cells in time has become a powerful approach to study cellular dynamics in cell lines and embryos[40]. Pioneering examples include tracking of embryos of *C. elegans*, fruit flies, zebrafish and mice[41-43]. In a recent paper, 10^4 cells were tracked in growing mouse embryos over 48 hours, during gastrulation and early organogenesis[44]. The authors imaged cell nuclei using adaptive multi-view light sheet imaging. Cells were tracked using a Gaussian mixture model, with the center of each nucleus

determined by fitting their fluorescent signal to a 3D Gaussian function. By assigning cells present at the end of the experiment to different tissues, e.g. the heart field or the neural tube, based on anatomical features, and by following the tracks of these cells backward in time, it was possible to reconstruct how these cells flowed out of the primitive streak and assembled into tissues. This revealed that both the timing and position of cells as they left the primitive streak was key to determining their cell fate. Moreover, by tracking cell divisions in time, the authors showed that the orientation of cell divisions changes several times during neural tube closure, with important impact on tissue morphology.

Despite the advance in analyzing mouse embryogenesis, the McDole study also underscores the formidable technical challenges that remain when studying development, including in organoids. While the tracking error rate was low enough to reconstruct the general flow of cells, it was too high to automatically reconstruct lineages in most parts of the embryo. This is because a single erroneous switch in cell identity can corrupt large parts of the lineage tree. Instead, the authors used a mosaic Cre/loxP reporter to sparsely label a small subsets of cells. This strongly reduced cell identity mistakes, by increasing the spacing between tracked cells, but severely reduced the lineage information that was captured. While such analysis of cell flow coupled to fate is useful for many applications, acquiring more exhaustive lineage information is of particular importance for organoids, for instance to identify the rare differentiation events and correlations between them, or to reveal spatial interactions on short length scales, including those between neighboring cells that originate from cell-cell signaling.

Following each cell without error requires both fine-tuned image analysis algorithms to accurately identify all nuclei and their positions, as well as a careful balance between limiting phototoxicity and increasing temporal and spatial imaging resolution. Similar to the embryo systems discussed above, organoids have extended 3D structures that lead to out-of-focus light and scattering, resulting in decreased resolution. In addition, cell nuclei tend to be more closely packed compared to early embryos and, particularly in epithelia, nuclei move rapidly along the apicobasal axis during division[45]. These properties require comparatively fast imaging (one 3D image every 5-15 min) at high spatial resolution (better than 1 μm/pixel), and generally complicates identification of all nuclei in each frame, for instance using gaussian fitting, and linking them through time without error. Light-sheet imaging may be used[39,46], which can limit resulting phototoxicity. This technique has also been used to study the flow of embryonic renal cells in kidney organoids during kidney rudiment re-aggregation[47]. However, more broadly available confocal imaging is often sufficient for organoid time-lapse imaging studies.

Apart from imaging, the dense 3D tissues found in organoids also pose challenges for nuclei identification using established image analysis approaches, such as Gaussian mixture models. An important recent advance in this regard is the use of neural networks and machine learning. This approach, which is based on a training procedure that uses manually analyzed

datasets to learn to identify nuclei, was shown to improve performance in non-organoid systems with closely packed cells[34,48]. Another issue is that cell tracking software can be difficult to use for non-experts[49]. Cell trackers often need to be reprogrammed, reconfigured, or retrained in the case of a neural network approach, upon changes in the system studied, imaging parameters, or fluorescent reporters, although algorithms that work for a wider range of microscopy images are developed [50]. Finally, an important practical problem is that software packages are not always well suited to correct for the tracking errors they invariably generate. This feature is less important when studying properties such as cell flow, but is important for lineage analysis in organoids, where differentiation events may be strongly influenced by stochasticity or neighbor interactions.

As a consequence, manual cell tracking approaches are still used, even for systems with hundreds of cells[51]. A promising new direction for organoid systems is to combine automated tracking based on neural networks with manual error correction steps. Such an approach was used to reconstruct cell lineages by tracking 50 cells during embryonic brain regionalization in brain organoids[52]. Key to scaling up such a hybrid approach from a limited number of lineages to entire organoids is to incorporate algorithms that automatically identify possible errors and allow for efficient manual correction of these errors. Recently, we developed such a hybrid approach to perform lineage tracking for whole intestinal organoids[53].

In the future, we expect that automatic cell tracking approaches will continue to improve, driven in part by advances in machine learning methods[40]. Currently cell tracking studies focus primarily on cellular movement and divisions. With automatic cell tracking becoming more accessible, a range of new applications will open up in organoid research, including the study of cellular differentiation, tissue renewal, shape and symmetry changes, and may involve simultaneous measurements of key regulatory and metabolic signals[54,55].

2.3 Cell tracking combined with fluorescent markers

The tracking of cells and their corresponding lineage trees by itself is often not enough to understand how developmental decisions are made. The results of these decisions, the cell fates, are invisible in most situations – with the exception of well-characterized systems where the cell type can be deduced from its spatiotemporal position and anatomical features, like the mouse embryo discussed in the previous section. By combining fluorescent markers that report on cell type with image-based cell tracking, it is possible to monitor a cell's identity, position and lineage dynamics concurrently, and hence study where and when cell fate decisions are made.

A recent study[56] showcases the power of this combination. This work focused on tissue regeneration upon damage of the neuromast, a small sensory organ, in zebrafish. An elegant combination of cytosolic and nuclear markers allowed imaging and identification of all three major cell types in the neuromast using only two colors. These fluorescence markers made it

possible to selectively photo-ablate different parts of the organ, and to subsequently study the regenerative potential of the different cell types. By manually tracking lineages it was found that multiple cell types have regenerative potential, but only one type has the potential to regenerate all three major cell types in the neuromast.

The authors then aimed to understand how these cell fate decisions are regulated to faithfully regenerate the organ from a single cell type. To analyze the large amounts of data generated during live imaging, feature lists were compiled for every tracked cell, including both intrinsic (e.g. the time since birth) and extrinsic information (absolute position, relative position to other cell types, polar orientation). Using a machine learning technique called 'random forest' to predict cell fate decisions, spatial features, like the position of cells relative to the organ center during division, were shown to be highly predictive of the cell fates that their progeny will take on. Intrinsic features were uninformative, suggesting that in this system, cell fates are not determined by (prior) cellular heterogeneity but by the cells plastically responding to their environment. This influence of position on cell fate would have been difficult to determine without the combination of live fluorescence markers and image-based lineage tracking.

For organoids, this approach is increasingly feasible, especially given recent progress in CRISPR based techniques that allow fluorescent reporters to be directly incorporated in organoid lines[32]. An organoid model for breast cancer has been used to study why some cells carrying an oncogenic mutation become highly proliferative while others do not[57]. Tracked lineages and a fluorescent reporter for the mutation indicated that the local density of mutated cells was the most predictive feature. Being within a cluster of other mutated cells yielded increased progeny. Again, the combination of spatial and lineage information (in this case the amount of progeny) provided by cell tracking were central to the conclusions.

We note that the breast organoids from Alladin et al. and the neuromasts studied by Viader-Llargues et al. are comparatively small systems and contain few cell types. Tracking will be more challenging in larger systems of several hundreds of cells, while spectral overlap limits the number of fluorescent labels, and hence the ability to distinguish all cell types of interest. Besides reporting for cell type, fluorescent proteins can also quantify cellular processes in organoids, such as chromosome and tubulin dynamics during cell division[32,45,58]. Even metabolic processes like oxygenation can be followed using fluorescence sensors[59]. Also promising are fluorescence reporters for the signaling pathways that regulate developmental decisions. Often, these reporters can be fused either directly to a downstream target of the pathway or placed under control of a target gene promotor. Short-lived fluorescent proteins might be required to detect rapid pathway activity dynamics[60]. Indeed, newly developed Wnt and Notch reporters have shown notable dynamics, which in turn impact differentiation[61-64]. These functional read-outs can be readily combined with lineage tracking to quantitatively study correlations with cellular organization and differentiation.

2.4 Cell tracking combined with end-point measurements

As discussed in the previous section, signaling dynamics and changes in cell state during development can in principle be monitored directly using fluorescent markers. However, the spectral overlap of fluorescent proteins and the time investments associated with the required genetic engineering limits the number of colors that can be imaged simultaneously to at most 2-3. This severely limits the number of cell types or genes that can be tracked in one experiment.

An approach that may help to circumvent this limitation was used to study 2D cultures of mouse embryonic stem cells[65]. After tracking the cells over time, these same cells were subsequently studied using three-color single molecule RNA-fluorescence in situ hybridization (RNA-FISH). These data quantified the expression of three genes that mark various differentiation stages. Each cell was classified as high or low for each gene, resulting in eight possible cell states. These states were correlated with the lineage history of each cell using an analytical approach called Kin Correlation Analysis (KCA). This method infers the cell-state transition rates during time-lapse imaging by analyzing state-correlations between relative cells, such as sisters and cousins. Reversible transitions occurred only between adjacent cell states, in a linear chain of cell states from pluripotent to more differentiated. Overall, these experiments show how developmental dynamics can be inferred by combining dynamic cell tracking with static end-point measurements.

One advantage of such end-point measurements is their scalability. In order to find genes important to E. coli cell cycle control, cells from a library of 235 CRISPR interference (CRISPRi) perturbations were tracked for days jointly in a single experiment, while characterizing phenotypes such as chromosome replication forks, cell size, and growth rates. Afterwards, ten sequential rounds of FISH labelling were performed to identify the underlying CRISPR perturbation in each tracked lineage. Key replication initiation regulators could hence be identified, and yielded new replication initiation control models[66]. Measurements of gene expression in single fibroblast cells and in brain tissue have been performed using a sequential staining method termed multiplexed error-robust FISH (MERFISH)[67-69]. Applying smFISH to organoids is challenging, due to limited penetration of probes and high background fluorescence, and often requires cryosectioning, which is non-trivial to implement[70]. However, a protocol for 3D smFISH in whole-mount colon organoids, that relied on reducing the background fluorescence of Matrigel, was described recently and may be applicable to other organoid systems[71].

This approach may be extended to other end-point analysis techniques such as immunostaining. Immunostaining has been often used to study cell type and differential expression statically, without a combination with cell tracking, for instance recently in neural tube development[72]. Multiplexed immunofluorescence imaging techniques allows for more than ten sequential rounds of antibody staining in a tissue[73]. Multiple rounds of immunostaining has been used in intestinal organoids[46], though not in combination with cell

tracking. An exciting idea is performing single-cell RNA-sequencing (scRNA-seq) as end-point analysis after cell tracking, as it would allow one to correlate the lineage tree to genome-wide changes in gene expression. However, as cells may then have to be removed from their tracked location to perform sequencing, which may be partially achieved by sorting fluorescently labelled cells or microaspiration, challenges remain in linking these different data sets. Alternatively, cell tracking could be followed by the recently developed spatial transcriptomics technique, which allows for RNA-sequencing while keeping the information of relative cell positions, though current techniques use sectioning and do not provide 3D information[74-76]. Overall, combining lineage tracking with static end-point measurements in organoids provides a promising approach to study spatio-temporal organization in developing organoids.

2.5 Discussion

In this review we have highlighted a number of recent studies that show the power of cell tracking approaches and how they can be applied to organoid model systems. They first of all underscore a general reality: the more we follow biological processes in time, the more crucial dynamics we uncover. This is perhaps particularly true for developmental processes, where organization in time is the core of the problem. They also illustrate how key technologies are now converging. On the one hand, advances in 3D microscopy and image analysis algorithms provide increasingly detailed views of cellular dynamics. On the other hand, rapid progress in genetic engineering and single cell sequencing yield ever more information on key regulators and markers of cell identity. With adjustments, these techniques are highly suited for application to the biology and biophysics of organoids, where study of temporal dynamics is still in its infancy. Together, these developments now provide an exciting opportunity to understand the underlying principles by which organs and organoids are organized in space and time.

While the studies reviewed here show that the main technologies are available, numerous improvements and extensions can be envisioned. At a practical level, more reliable automated tracking of cell movement and division would greatly expand the general use and throughput of cell tracking approaches. The development of new fluorescent reporters and sensors promises far more detailed observations of regulatory and metabolic pathways than is possible currently. Approaches that can link cell tracking to more expansive cell expression measurements, including multiplexed immunostaining, smFISH, and single-cell sequencing, have the potential to unlock a next level of understanding down to the molecular scale.

A general challenge will be to analyze the resulting lineage tracking data sets, given their complex, multi-faceted nature that combines space, time, lineage and internal states. How can insight into simple organizing principles be inferred from such data sets? One approach could be to use machine learning techniques to identify the most relevant features informing cell fate decisions[56,57,76]. In addition, one may exploit dimensionality reduction techniques

similar to ones used in the single cell sequencing and flow cytometry fields, which deal with similar highly dimensional single-cell resolved data, as they could identify hidden structures in the data. For instance, currently expressed cellular fate could be determined by past cell-cell contacts, cellular location, orientation, and molecular signals, while conversely, cellular migration speeds and spatial patterns may depend on cell type, age, and genealogical relations. Reduction techniques are already being used to analyze high throughput microscopy[77,78] and recently have been used in the analysis of time-resolved imaging of whole organoids[46] and single cells in 2D culture[79]. Beyond the single cell level, recently developed methods allow statistical analysis of lineage trees shapes[80,81] and simultaneously measured phenotypic signals[54,80,82-84]. In this way key differences between lineages may be identified and correlated with developmental decisions. However, it is likely that new analysis methods and the introduction of bottom-up mechanistic models are needed to make full use of the incredible richness of information that these new technologies can provide, and allow one to move from a descriptive to a predictive understanding of organ biology.

3 Cell tracking with accurate error prediction

Cell tracking is an indispensable tool for studying development by time-lapse imaging. However, existing cell trackers cannot assign confidence to predicted tracks, which prohibits fully automated analysis without manual curation. We present a fundamental advance: an algorithm that combines neural networks with statistical physics to determine cell tracks with error probabilities for each step in the track. From these we can obtain error probabilities for any tracking feature, from cell cycles to lineage trees, that function like p-values in data interpretation. Our method greatly speeds up tracking analysis by limiting manual curation to rare low-confidence tracking steps. Importantly, it also enables fully automated analysis by retaining only high-confidence track segments, which we demonstrate by analyzing cell cycles and differentiation events at scale, for thousands of cells in multiple intestinal organoids. Our approach brings cell dynamics-based organoid screening within reach and enables transparent reporting of cell tracking results and associated scientific claims.[1]

[1] The content of this chapter is published as: Betjes, M.A., Kok, R.N.U., Tans, S.J. *et al.* Cell tracking with accurate error prediction. *Nat Methods* 22, 2400–2410 (2025)

3.1 Introduction

Cell proliferation, differentiation, movement, and their organization in complex cell lineages are key to understanding organ homeostasis and associated diseases. The development of organoid cultures, which recapitulate key features of organ development *ex vivo*[85,86], has enabled the study of developmental dynamics at the single cell level using time lapse microscopy[8,53,87-90]. To address the complex challenge of analyzing the dynamics of hundreds of cells in dense 3D organoid architectures over multiple generations, AI-driven semi-automated algorithms have been developed that track cells based on their fluorescently labelled nuclei[53,87,91-94].

However, all current cell tracking approaches face a fundamental limitation: algorithms output a single tracking solution among many possible solutions, and are highly prone to making errors, yet lack a statistical basis to quantify prediction uncertainty (**figure 3.1A**). This lack of statistical interpretability makes rigorous analysis based on cell tracks impossible, as the inability to assess the confidence of cell tracking-based results can lead to unfounded conclusions and, more generally, limits scientific transparency and reproducibility. Finally, the black-box nature of cell tracking even hampers method development and optimization itself as it makes it difficult to identify and tackle the true source of tracking errors. In contrast, other widely-used bioinformatics methods like sequence alignment[95,96] or differential gene analysis[97], do inherently provide statistics on their output, with the resulting confidence in data interpretation and reporting crucial to their widespread adoption.

These problems are particularly acute in context of development and tissue homeostasis, where error in even a single tracking step can radically alter the biological interpretation (**figure 3.1A**). For organoids, additional tracking challenges are presented by closely-packed nuclei that move rapidly during cell division[45]. While the recent adoption of neural networks in cell tracking algorithms has greatly increased tracking quality[53,91,92], current methods are far from error-free, especially in organoids[53,87]. Existing methods use *ad hoc* heuristics, like rapid nuclear volume changes or large cell displacements, to flag potential errors for manual correction[53,87]. These methods rely on manually set cut-offs and user interpretation, which hampers reproducibility and assessment of data quality. Moreover, because such heuristics cannot provide any measure of confidence for the obtained cell tracks, creating error-free datasets relies on extensive manual curation, up to the point of checking essentially each tracking step. This process can take days for a single 300-500 cell organoid, which burdens users and limits tracking applications. In particular, screening different growth conditions or mutant backgrounds is prohibitively time consuming to analyze.

Here, we present a conceptually new approach: an algorithm that determines both cell trajectories and their error rates (**figure 3.1B**). Building on our previously developed OrganoidTracker[53], we introduced two major innovations: first, we show that neural networks can perform key tracking tasks while outputting ‘calibrated’ likelihoods, meaning

they provide accurate estimates of error probability of their prediction. This enables splitting the tracking procedure into specific sub-tasks that can be independently optimized, while their calibrated probabilities can be combined to give a single, global probability estimate. We developed dedicated neural networks for the sub-tasks of linking cells between subsequent microscopy images and that of identifying division events. Second, we used concepts from statistical physics, including microstates, partition functions and marginalization, to compute 'context aware' error probabilities. In practice, this procedure implements our intuition that a low-probability tracking step can in fact be of high-confidence, if all alternative cell linking arrangements are excluded by high-confidence tracks of surrounding cells.

Importantly, these innovations now also enable the reporting of statistical significance. The resulting OrganoidTracker 2.0 provides error probabilities for any lineage feature of interest, from cell cycles to entire lineage trees. These error probabilities can then be used to asses and report the statistical significance of conclusions based on these tracking features, performing a role similar to p-values. Our innovations also enhanced tracking performance. First, OrganoidTracker 2.0 is a highly competitive cell tracker, with output tracks containing errors at <0.5%/cell/frame for intestinal organoid data, even before manual curation. Moreover, it drastically sped up this manual curation, by focusing it on those parts of cell tracks that had high predicted error rates. A 60 hour movie with over 300 cells tracked for over 300 time points was curated in ~3 hours rather than 3 days. Second, the resulting method enables fully automated analysis without any human curation, by removing, instead of reviewing, the low-confidence parts of cell tracks and using the high-confidence parts for further analysis. Demonstrating the power of this approach, we extracted cell cycle time and proliferation rates for 20 organoids in fully automated manner, thus opening up the possibility of high-throughput screening of cellular dynamics. We show that OrganoidTracker 2.0 also provides excellent automated tracking for mouse blastocysts and C. *elegans* embryos, with its performance for the latter ranking as the best performing tracking algorithm on the Cell Tracking Challenge[98]. Furthermore, we provide an easy user interface, extensive documentation and straightforward retraining procedures for different biological model systems. Our framework is modular and can thus readily incorporate other cell tracking functions, such as automated cell death detection. Finally, the statistical basis provided by our method will be important for data sharing and statistically justifying scientific claims.

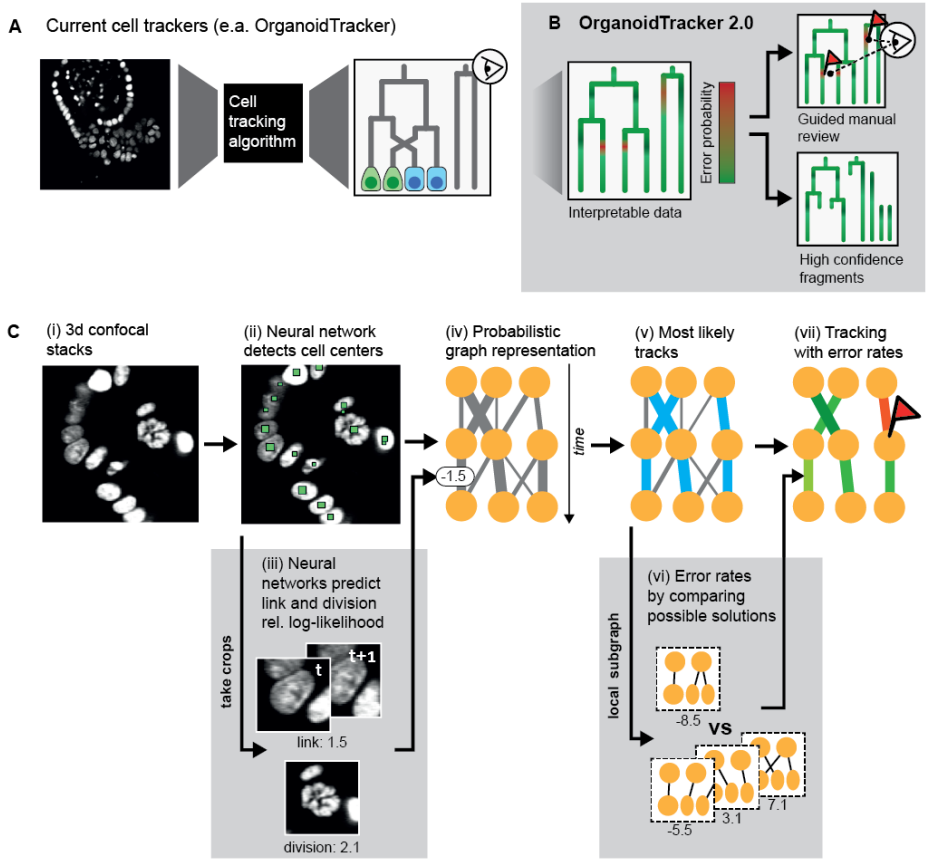


Figure 3.1: **Method overview. A)** Current cell tracking algorithms convert microscopy images into cell tracks without providing information on accuracy. Yet even single errors can greatly alter the biological interpretation of lineages (here, change in symmetry of divisions). Hence, extensive manual review is required and finally no assessment of statistical confidence can be provided. **B)** OrganoidTracker 2.0 outputs not only tracks but also associated error rate estimates, greatly aiding data interpretability and transparency. These error estimates also enable a drastically reduced manual review, or fully automated filtering to achieve high confidence datasets. **C)** Method workflow, highlighting two novel components (grey boxes). (i) Generation of 3D confocal stacks of nuclear marker fluorescence. (ii) Neural network detection of nuclei centers. (iii) Neural network prediction of cell linking or division probabilities, based on image crops. (iv) Constructing a graph representation of the tracking problem, based on predicted link and division probabilities. (v) Determination of the globally optimal solution representing the most likely cell trajectories. (vi) Estimating link error rates through systematic comparison with alternative tracking solutions. (vii) Predicted cell tracks with error rate predictions for individual links

3.2 Results

3.2.1 Method overview

Our method is divided into two parts: first, we use neural networks to identify the cells in each frame and predict the probabilities of all possible links between them (**figure 3.2C**, i-iv). Next, we use these results to find the most likely tracks and compute their associated error rates (**figure 1C,** v-vii). Central to our approach is the concept of a probabilistic graph description of the tracking problem[99] (**figure 3.2A**). Here, each node represents a cell detected at one time point, as obtained by a neural network that predicts cell center locations from 3D images of fluorescently labeled nuclei. Links between the nodes represent possible connections between these cell detections, through movement or division. To each link, we assign a 'link energy', defined as the negative relative log-likelihood of a link being true based on the neural network prediction, so that a low energy indicates a more plausible link. Similarly, we determine a 'division energy' for each node that indicates the likelihood of division. Expressing predictions of likelihoods as energies allows us to use concepts from statistical physics to analyze and combine these predictions. A key innovation is that we employ neural networks to predict these link and division likelihoods based on microscopy data. Here, we leverage the property of neural networks that their output scores can function as accurate probability estimates[100], which has thus far rarely been used in tracking applications. Using an integer flow solver, we can find the collection of paths on the graph with the minimal associated energy[99], corresponding to the most probable set of cell trajectories. Finally, we use the link energies and graph structure to compute error probabilities for every link in the predicted tracks, thereby providing us with both the trajectory predictions and their associated error rates. Below, we discuss each step in more detail.

3.2.2 Cell detection

We detect cell centers using a 3D-UNet neural network[101] (**figure 3.2C**). Specifically, this network uses 3D images of organoids carrying a fluorescent nuclear marker (H2B-mCherry, **supplementary figure 3.1**) to predict a distance map, which for every pixel records its distance to the closest cell center[53,92]. Cell centers are then obtained by finding local peaks in this distance map. This approach enables the generation of training data by annotating cell centers, which is less labor-intensive than manual 3D segmentation of nuclei. An inherent challenge to distance maps is that cell center peaks for closely packed nuclei blend into one another, causing under-segmentation. We therefore developed an adaptive distance map, in which we assign increased distance values to pixels that are almost equidistant to two cell centers (**figure 3.2B, supplementary figure 3.2A**). This ensured that cells remain well-separated in the resulting map, thus reducing segmentation errors (**supplementary figure 3.2B**). To test the impact of the adaptive distance map on cell center detection (together with additional improvements in the training data generation and augmentation pipeline, see

Methods), we compared performance to OrganoidTracker 1.0. Both methods were trained on the same data set and subsequently tested on identical unfamiliar microscopy images. Overall, detection accuracy was improved substantially, decreasing the error rate ~4-fold over the complete time-lapse data set. The high accuracy only decreased slightly, from 99% to 95%, when the cell's nuclear fluorescence had poor signal-to-noise ratio, for example after prolonged imaging (>50 hours, **figure 3.1D**) or deep in the imaging volume (>40 μm, **figure 3.1E**). Finally, predicted cell centroids closely aligned with the center of mass of each cell's 3D nuclear shape (**supplementary figure 3.3**).

3.2.3 Estimating link and division probability

We then construct the linking graph by connecting each node, representing a detected cell, with all possible links. For practical purposes, we culled links representing unrealistically large displacements (**Methods**). Links either connect the same cell in two consecutive frames or connect a mother and daughter cell. The first category of cells only move and generally retain similar nuclear shape and fluorescence intensity between frames, while dividing cells show distinct changes in nuclear morphology. To deal with this heterogeneous data, we trained a neural network to estimate the relative log-likelihood of a link being correct, based directly on the imaging data. Here, we employ a fundamental ability of classification neural networks that use a cross-entropy loss during training, especially when combined with Platt scaling subsequently[100,102], namely that the likelihood they predict exactly matches the actual likelihood of the event being correct as observed in ground-truth data (referred to as a 'well-calibrated likelihood'). This property of neural networks has thus far not been used in cell tracking and is key for our error rate estimation.

We designed a neural network that inputs cropped 3D images centered on each cell's detected position for time points t and $t+1$ and predicts the likelihood that they represent the same cell, using information within the crop on both the nucleus and its local environment (**figure 3.2F**). The trained network correctly assigned low energy (high likelihood) to links between the same cell, even when the fluorescence signal changed substantially, while assigning high energy (low likelihood) for links connecting a cell to its neighbor, with 'energy' defined as the negative relative log-likelihood. We compared the network's performance to the baseline criterion, often used for tracking[103], that the links representing the smallest displacement between frames are correct. First, when comparing all possible links in the graph against ground-truth data, we found that links corresponding to small displacements (<3 μm) were often correct (**figure 3.2G**), while large displacement links (>7 μm) were incorrect. However, many correct links represented substantially larger displacement (3-7 μm). These large displacements occurred predominantly during cell division (**figure 3.2G**, inset, **supplementary figure 3.4**), and their correct assignment is thus essential for lineage tree reconstruction. We therefore compared the accuracy of links assigned by the neural network (link considered true when predicted probability>50%) to those assigned by the 'smallest

displacement' criterion (for each cell, only the shortest-distance link towards it is true) as function of link displacement. For <3µm displacement, both approaches yielded similar high accuracy (**figure 3.2H**). However, the accuracy decreased strongly for larger displacements when using the 'smallest displacement' criterion, but remained high for the neural network, also for cells close to division (**supplementary figure 3.4A-D**) and fast-moving non-dividing cells (**supplementary figure 3.5A,B**), demonstrating the superiority of our neural network approach.

While most graph nodes have a single correct outgoing link, reflecting movement, nodes of dividing cells have two correct links connecting to its two daughters. Dividing cells exhibit distinct nuclear morphology, with chromosomes forming the metaphase plate, which we exploited to determine the likelihood that a node corresponds to a dividing cell and thus allowed to have two outgoing links. We designed an additional neural network that used 3D image crops to predict division likelihood, including the previous and subsequent frame to facilitate identifying the exact moment of division (**figure 3.2I**). When testing images at different times relative to division, defined as the last frame before chromosome separation, division assignment (>50% probability) indeed coincided with the moment of division in >90% of cases (**figure 3.3J**). Moreover, cells at time points before or after division were only rarely assigned as dividing, even though they are visually similar to cells at the exact time of division.

We also substantially improved prediction by up-sampling challenging cases during training: fast-moving, dividing cells for link prediction (**supplementary figure 3.6**) and time points around division and cell death for division prediction (**supplementary figure 3.7**), as dying cells resembled dividing cells in appearance. This is done in practice by presenting these cases more frequently to the networks during training than their natural occurrence. The general ability to tailor training data sets to individual tasks (cell detection, link and division detection) is a major advantage of our modular approach, compared to merging multiple tasks in a single, more complex neural network[87,92]. This fine-tuning also shows the benefits of having directly interpretable probabilities, instead of generic weights[53,92], as the output of tracking subroutines in guiding method optimization. Finally, we validated that neural network output indeed represented true probabilities. We binned all possible links based on their predicted likelihood to be correct. Then for each bin, we calculated the true likelihood, i.e. the fraction of links that were correct according to the ground-truth data. We found that predicted likelihoods were well-calibrated, with predicted and true link likelihood matching for the full likelihood range (**figure 3.2B**). This ability to accurately predict likelihoods is essential for estimating tracking errors, as discussed below.

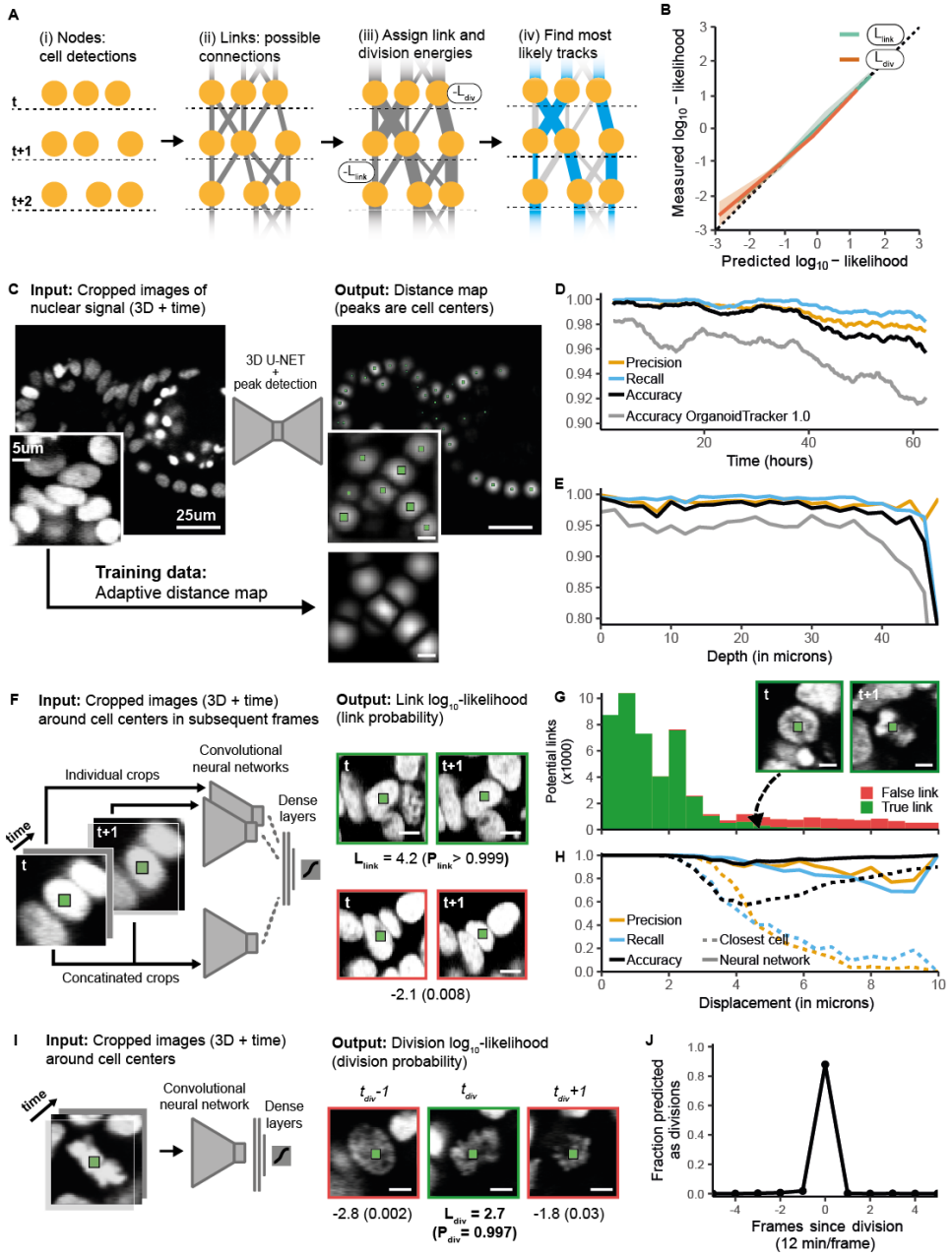


Figure 3.2: Probabilistic graph construction by neural networks. (A) Probabilistic graph workflow. Nodes are detected cells, and grey lines are possible links that connect cells between subsequent time points, reflecting either movement or division. Thicker lines indicate more likely links, or a lower 'energy'. Blue lines indicate the globally optimal solution. Cell detection (i) and link and division likelihood prediction (iii) is performed by neural networks. **(B)** The predicted relative log-likelihoods by the neural networks strongly correlate with measured relative log-likelihoods (the probability of being true in the manually annotated control), both for links and divisions. The dotted line corresponds to perfect calibration. Data represents n=5 organoids, with shaded region denoting standard

deviation around the mean. **(C)** 3D-UNET neural network trained to generate a distance map that indicates proximity to nuclei centers. Cell centers (green squares) are obtained by peak finding. Smaller squares indicate cell centers located below or above the z-slice shown. **(D), (E)** Accuracy of cell detection, compared to OrganoidTracker 1.0, as function of (C) time and (D) imaging depth, for one organoid data set. Metrics are averaged over 10 frames. Only cells <40 um deep were included for (C). **(F)** Convolutional neural network trained to predict link likelihoods, based on crops centered around two cells detected in subsequent time points. Output images demonstrate a high (green) and low (red) likelihood link prediction, corresponding to a true and false link, respectively. **(G)** Link analysis for manually curated data. Links representing small cell displacements are typically true (green) and large-displacement links false (red). Both true and false links are observed in between, demonstrating that displacement alone does not determine link correctness. Images show a correct large-displacement link, of a cell undergoing division. (**H)** Link prediction accuracy. For all displacements, the neural network strongly outperforms predictions based on the "smallest displacement" criterion, which assigns the link that minimizes displacement as correct. **(I)** Neural network trained to predict division likelihood based on image crops centered at detected cells. Images show three subsequent frames of a dividing cell: just before chromosome separation (green border, high predicted likelihood), and before and after division (red, low predicted likelihood). **(J)** Fraction of cell crops assigned as dividing (>50% probability) versus time relative to division, defined as the last time point before chromosome segregation. Division assignments occur predominantly at the exact measured division time.

3.2.4 Track prediction

To construct cell trajectories, we use a min cost flow solver algorithm[99] to select the set of links in the probabilistic graph that globally minimize energy, and thus maximizes the probability of the tracking solution. While close to optimal tracks are obtained readily, the algorithm does not guarantee identification of the global optimum and we typically found minor mistakes, such as link pairs that decrease global energy when swapped. Moreover, flow solvers cannot change the graph structure by adding or merging nodes, causing vulnerability to under- and over-segmentation[104]. We therefore automatically check if overall probability is increased by swapping link pairs and by adding or removing nodes (**Methods**). In the latter case, the algorithm identifies pairs of disconnected tracks that can be joined by adding a node, corresponding to a missed cell detection, or by merging nodes, corresponding to over-segmentation (**supplementary figure 3.8A,B**). New links and nodes are assigned energies reflecting the probability of under- and over-segmentation, as given by the detection network's recall and precision. These steps are performed automatically and with the required parameters obtained from the network predictions, without any further user input needed. Overall, this procedure substantially increases the duration over which cells can be continuously tracked and shows the benefits of our probabilistic framework in data post-processing (**supplementary figure 3.8C**).

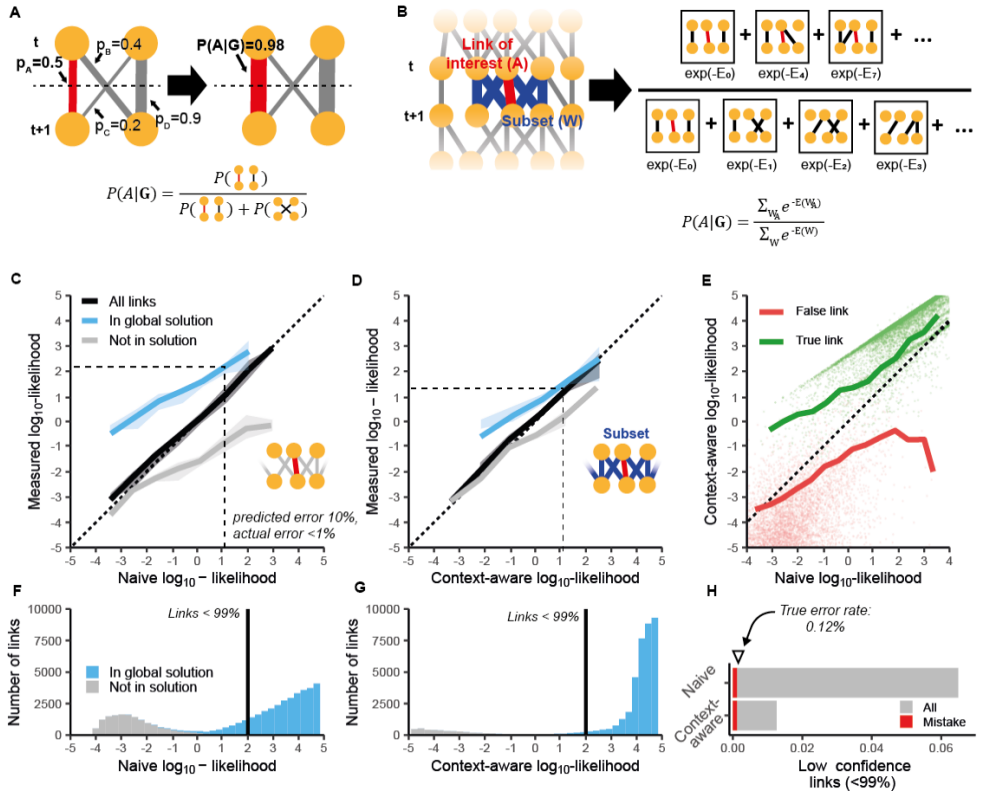


Figure 3.3: Error rate estimation by marginalization. (A) Simplified example explaining marginalization. Lines indicate putative links *A-D*, with thickness indicating their estimated probability. Link *A* (red) has low a low predicted probability. However, as links *B* and *C* are unlikely because *D* has a high probability of being true, it implies *A* is true with high certainty. The probability P(A|G) that *A* is true given graph structure *G,* can be calculated by looking at the probability of configurations containing link A versus ones that do not. **(B)** Schematic outline of marginalization, performed on a subset of links around the link of interest. P(A|G) is given by the summed energy of all configurations containing link *A* normalized by the summed energy of all configurations. **(C), (D)** Measured link likelihood versus (C) naive likelihoods predicted by the neural network or (D) context-aware likelihoods calculated by marginalization. Data is shown for all possible links (black) or links that are either in the global solution (blue) or not (grey). For naive likelihoods (C), links in the tracking solution are more likely correct than expected on the basis of the neural network prediction, while for context-aware likelihoods (D) they more closely match measured likelihoods, reflecting integration of graph information. Dotted line is perfect calibration. Data for n=5 organoids. Shaded region is S.D.M. **(E)** Context-aware likelihoods versus naive likelihoods. Dots are individual links and lines are averages, for true (green) or false links (red). Marginalization increased predicted likelihood of correct links while decreasing it for incorrect links. **(F), (G)** Number of links versus predicted (F) naive or (G) context-aware link likelihood. While for (F) most links in the globally optimal solution (blue) are predicted as high-confidence (>99% probability), a fraction have confidence levels similar to rejected links (grey). In contrast, for (G) virtually all globally optimal links are now predicted as high-confidence. **(H)** Fraction of links in the globally optimal solution deemed low-confidence (<99% likelihood) after marginalization. Red is the fraction of low-confidence links that were actual errors compared to ground-truth, while the arrow indicates the fraction of errors in all links. Almost all errors that occurred were covered by the <99% probability threshold. Marginalization thus reclassified many low-confidence links as high-confidence, but not those that represent errors.

3.2.5 Context-aware estimation of link error

Central to our approach is estimating the error rate of individual links. The 'naive' link likelihoods, as predicted by the neural network, in principle provide information on each link's error probability, but ignore the context of the link predictions made for surrounding cells. The importance of context is evident already in manual tracking: here, human trackers typically first establish high-confidence links, which in turn, by reducing the remaining possible links, facilitates subsequent assignment of lower-confidence links. Such contextual information can be computed in our probabilistic framework as well. This is illustrated by the simplified graph in **Figure 3A**, where the likelihood of a low-confidence individual link being true is increased dramatically (from 50% to 98%) in the context of the larger graph, because high-likelihood links exclude all alternative linking arrangements. To generalize this notion, we considered the combined likelihood of all possible linking arrangements containing the link of interest based on the neural network predictions, using the following approach (**Figure 3B**): for each possible linking arrangement, corresponding to a tracking solution W_i, we calculate its energy $E(W_i)$ by summing all link and division energies, with the solution likelihood proportional to $e^{-E(W_i)}$. The 'context-aware' likelihood of link A is then given by the total likelihood of all tracking solutions containing that link, $\sum e^{-E(W_{i,A})}$, normalized by the sum for all possible solutions, $\sum e^{-E(W_i)}$. We call this procedure marginalization, as all other variables are marginalized out to arrive at a single-link error rate estimate without referencing any other links. Since summing over all possible tracking solutions is computationally unfeasible, we restricted this to local subgraphs of links less than 3 steps away (**figure 3.3B**, **Methods**). Increasing subgraphs to 4 links away did not significantly improve prediction accuracy (**supplementary figure 3.7, Methods).** For 3-link subgraphs, marginalization required <1 hour for a 60 hour time-lapse data set.

Our marginalization procedure assumes that the individual link and division predictions are independent. This is not strictly true, as these predictions are partially based on shared inputs: for example, the various image crops used as input can overlap. Assuming predictions represent independent evidence causes overconfident error predictions when combined, meaning that links deemed very likely (low negative energy) are more often false, and links deemed unlikely (high positive energy) more often true than predicted (**supplementary figure 3.9**). To correct this, we employed the similarity with statistical physics to introduce a 'temperature' T that decreases energies to E_i/T for every neural network prediction i. For T>1, this reduces the confidence of individual predictions to compensate for the overconfidence introduced during the marginalization procedure. We obtain the optimal value of T by calibrating the marginalized predictions against the ground truth. For this, the same data sets already used for neural network training and validation are used, without any further user input required (**supplementary figure 3.9, Methods**).

Overall, our marginalization approach borrows conceptually from statistical physics, with each possible tracking solution equivalent to a microstate and the normalization factor to the

partition function. From a probabilistic perspective, our method resembles the multiplicative opinion pooling framework[105,106], where different opinions (here, neural network predictions) are combined assuming that their pooled probability is proportional to the product of probabilities associated with the individual opinions. Here, we have extended this framework to predictions that concern different events (link and divisions) that are part of a larger structure (the graph), instead of simply reflecting the same event (see **Methods & Appendix).**

3.2.6 Evaluation of error rate predictions

We compared both naive and context-aware error rate predictions against measured error rates, by testing them against manually annotated data sets. To avoid bias, these datasets were generated completely independently from the OrganoidTracker pipeline. We then reused the ground truth cell centers to generate the link predictions and calculated the context aware error rates. First, we examined naive link predictions, i.e. before marginalization. Overall, naïve predicted likelihoods were well-calibrated, correlating strongly with measured likelihood (**figure 3.3C**). However, our data indeed indicated that the error predictions are under-confident in determining true and false links without access to contextual tracking information. Specifically, links identified by the flow-solver as part of the globally optimal solution displayed measured likelihoods much higher than naive neural network predictions, while for links rejected from the global solution measured likelihoods were lower than predicted (**figure 3.3C**). This matches our intuition that the graph contains additional information on link likelihood (**figure 3.3A**), as the flow-solver selects links based on the complete graph information, while the neural network makes its naive predictions based only on local image information. In contrast, context-aware link predictions showed strongly improved confidence, reducing the mismatch between predicted and measured likelihood for both flow-solver selected links and the rejected links (**figure 3.3D**). Moreover, incorporating graph context specifically increased the predicted likelihood of true links, while decreasing it for false links (**figure 3.3E**).

The improved context-aware link predictions have significant practical advantages for error correction. For naive predictions, a large fraction (6%) of links selected by the flow-solver must be reviewed, when using <99% predicted probability as threshold for manual curation. At the same time there was significant overlap in likelihood values between the flow-solver selected links and those it rejected (**figure 3.3F**), erroneously suggesting that these links were equally likely to be true. However, for context-aware predictions, the predicted likelihoods of flow-solver selected links differed strongly from rejected links (**figure 3.3G**). Moreover, the number of <99% confidence links that must be manually reviewed reduced substantially (1%), while practically all true linking mistakes (0.12%) were still detected (**figure 3H**). An order of magnitude more links (~25%) must be reviewed to achieve similar accuracy using a simple heuristic based on cell displacement (**supplementary figure 3.10**). Finally, we note

that our marginalization procedure specifically benefits otherwise difficult to predict links representing large cell displacements (**supplementary figure 3.5C)**.

As a final control we tested the marginalization procedure in the context of our full pipeline. We ran our algorithm on three different organoids. We then performed manual curation to generate ground truth data to which the predicted error rates could be compared. Again, we found that the error rates were well-calibrated (**supplementary figure 3.11)**. Lastly, we could reduce the computation time without reducing accuracy by excluding highly unlikely links before marginalization, as this did not significantly impact error rate computation (**supplementary figure 3.11, Methods**).

3.2.7 High-level error probabilities for lineage features and manual curation

We ran our pipeline on a ~60 hour time-lapse dataset of a representative organoid (**supplementary videos 1, 2 & 3)**. Marginalization yielded cell tracks and lineages with an error rate for each link (**figure 3.4A**). The rate of predicted potential errors (defined as <99% probability links) was 1.5%/cell/frame, after removing tracks deep in the imaging volume, with potential errors predominantly, but not exclusively, concerning divisions (**figure 3.4A)**. These error rates can now be used to propagate the error probabilities to downstream complex lineage features as $p=1-\prod_i P_i$, where P_i are the context-aware probabilities of all links *i* of the lineage feature of interest, with p representing the probability that the feature is not correctly tracked (**supplementary figure 3.12)**. These high-level error probabilities enable users to assess the statistical significance of, for example, individual cell cycles, the observation that cells are sisters, or even entire lineage trees (**figure 3.4A,B, Methods**). In this manner, these probabilities can function akin to p-values, although we note that they do not follow from a hypothesis testing framework. Moreover, we identified high-confidence lineage fragments as fragments with $p<0.01$ calculated over all links within the fragment. We found that long stretches of unprocessed tracking data were of such high-confidence, with high-confidence fragments often spanning multiple cell cycles (**figure 3.4B, supplementary figure 3.13**).

Our accurate knowledge of link probabilities implied that minimal manual correction focusing only on low-confidence links should be sufficient to obtain error-free tracks. To test this approach, we manually reviewed all links with <99% probability. We also reviewed all beginnings and endings of cell tracks mid-experiment (0.9%/cell/frame), which represented either cell death, cells entering or exiting the field of view or, potentially, cell detection errors. Of these potential errors, only a fraction represented true linking or detection errors (0.3%/cell/frame, **figure 3.4D**). However, correcting the few true errors did strongly improve the lineage trees, complementing them with previously unconnected sub-trees. (**figure 3.4C, supplementary figure 3.14**), underscoring the importance of identifying even infrequent tracking errors when complete lineage trees must be produced. Finally, independent manual tracking of cells in the lineages of **figure 3.4C** yielded identical trees. Correction required ~4

hours for a data set where ~300 distinct cells were tracked in a ~60 hour time window (**figure 3.4E, supplementary figure 3.14, supplementary video 4**). When we calculated error probabilities for cell lineages after manual correction, assigning a probability of one to manually corrected links and recomputing the marginalized link-likelihoods (**Methods**), we generally found low values, P<0.05 for all analyzed lineage trees (**figure 3.4C**), indicating high resulting confidence.

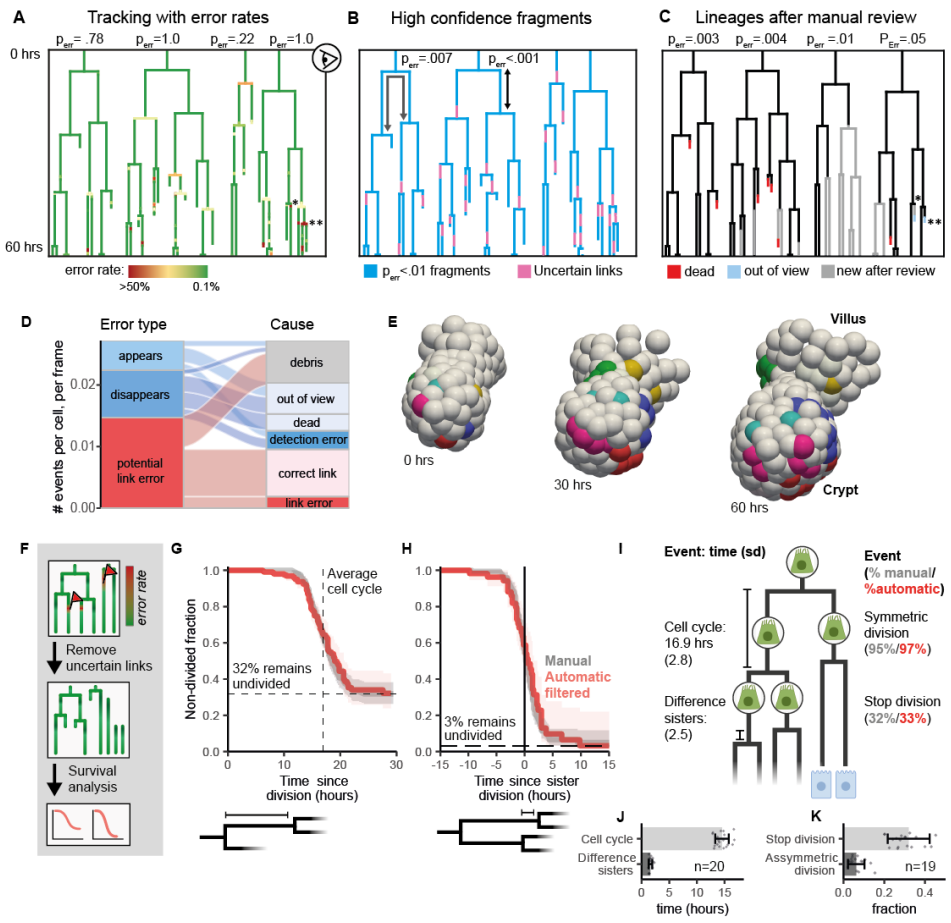


Figure 3.4: Applications. (A) Selected lineages before review, with color indicating associated error rates. Low-confidence links are yellow to red, (**) indicates a flagged error associated with a clearly erroneous lineage structure (unrealistically short cell cycle), while (*) would not be identifiable as an error based on lineage structure alone. Error probabilities are calculated for each lineage tree by combining error rates of all underlying links, identifying lineage trees as low-confidence even though they visually appear plausible. **(B)** High-confidence lineage fragments (<0.01 error rate) obtained based on individual link error rates. Users can identify high-confidence cell cycles (black arrow) or dividing cells that are sisters (grey arrow) without manual review. **(C)** Lineage trees after manual review of potential errors flagged in (A). Grey lines: lineage sections added as consequence of curating erroneous links. Error probabilities now indicate high confidence that lineages truly reflect underlying cell dynamics. **(D)** Characterization of potential errors. Links flagged as potential errors either represent (dis)appearing cells (blue) or are low-confidence (red). A significant proportion of potential errors represented very short tracks of cellular debris,

whose correction impacted the data only minimally. Only a minority of potential errors are actual errors in cell detection or linking that required correction. **(E)** 3D reconstructions including lineages in (C). Colors indicate cells in the same lineage. **(F)** Automated analysis without manual review, by filtering out low-confidence links and performing survival analysis on the resulting, partly censored data. **(G)** Survival curve indicating the probability that a cell has not divided at time *t* after its birth. Shown are manually annotated (grey) and automatically filtered data (red) for all cells in a single organoid. Vertical line denotes average cell cycle duration, while horizontal line shows inferred fraction of cells that stop dividing after birth. Shaded region is 95% confidence interval. **(H)** Survival function indicating the probability that a cell has not divided, as a function of the time relative to the division of its sister. Almost all cells divide within a 10 hour window around the time of sister division, showing that symmetric divisions dominate intestinal organoid growth. Shaded region: 95% confidence interval. **(I)** Comparison of key lineage dynamics parameters obtained by fully automated (red, no manual review) or manual analysis (grey). Fully automated analysis shows excellent agreement with manually annotated data. **(J)** Automatically obtained cell cycle times (grey) and differences between sisters (dark grey) for twenty organoids. Dots represent individual organoids and error bars the standard deviation. **(K)** Automatically obtained fractions of cells that stop dividing, for all cells (grey) and cells that have a dividing sister (dark grey) for nineteen organoids.

3.2.8 Fully automated lineage tracking by error filtering

Current 3D cell tracking algorithms require extensive manual curation after automated cell tracking before analysis is performed. Our ability to accurately estimate linking error rates enables a novel and fundamentally different approach: to remove low confidence track fragments and subsequently perform analysis only on the remaining high-confidence fragments, without any human oversight (**figure 3.4B and F**). For fragments that have high-confidence data from division to division, properties such as cell cycle durations can be directly measured and compared between different organoids and regions within an organoid (**figure 3.4B**). We also employed survival analysis (**Methods**), a statistical framework for dealing with censored (incomplete) data[107-109], to quantify a broad range of lineage properties while also incorporating information from lineage fragments containing incomplete cell cycles. Specifically, we generated Kaplan-Meier survival curves to estimate the fraction of non-divided cells as a function of time since cell birth (**figure 3.4G**), using all (incomplete) high-confidence track fragments that included at least one birth. This survival curve displayed two notable features: first, it plateaued at 32%, representing the fraction of cells that do not divide again and, hence, have differentiated. Second, the average cell cycle time is given by the time at which 50% of the dividing fraction have divided, with standard deviation corresponding to the spread around this time, yielding a cell cycle duration of 17±2.8 hour. We then extended this analysis to sisters, using sister pair fragments to generate survival curves as a function of time relative to sister cell division (**figure 3.4H**). Here, the curve plateaued at 3%, representing the fraction of sisters where one proliferated while the other ceased proliferation. This high degree of symmetry between sister cells is consistent with recent work[8,88]. Moreover, sisters had highly similar cell cycle durations, as indicated by the steep decline in the curve, corresponding to a difference in cell cycle duration of 2.5 hours between sisters. Overall, survival curves generated from automatically filtered and manually tracked data showed almost exact overlap (**figure 3.4H**). Finally, we demonstrated the automated nature of this approach by analyzing 20 different organoids (**figure 3.4J and**

K). We consistently found similar parameter values and survival curves, even as organoids displayed notable differences in size and morphology, indicating that the underlying lineage dynamics is independent of this morphological variation (**figure 3.4J and K, supplementary figure 3.15**).

3.2.9 Out-of-sample capabilities without neural network retraining

Next, we tested the performance of the organoid-trained neural networks on a diverse range of out-of-sample data. Performance could be impacted on two distinct levels. First, out-of-sample data can degrade cell detection and linking performance, leading to poor tracking. Second, even if tracking performance is not impacted, out-of-sample data can result in our algorithm predicting inaccurate error probabilities. These inaccuracies could then lead to less reliable manual curation or automated analysis.

We first examined performance on intestinal organoids exposed to the cell cycle inhibitor palbociclib. Visually, palbociclib exposure changed both cell appearance and dynamical behavior, with inhibition of cell division causing smaller nuclei and reduced overall cell movement. Nonetheless, cells could often be readily tracked through ~40 hour lineages without manual curation (**supplementary figure 3.16A**). The impact of palbociclib was evident by a comparative lack in cell number growth, while automated lineage analysis by error filtering demonstrated the expected strong inhibition of cell division (**supplementary figure 3.16B,C**). We next tested performance on organoids under normal conditions but imaged on a different scanning confocal microscope (**Methods**). Differences include lower pixel resolution, lower signal-to-noise ratio (SNR) and an objective with higher working distance, with the latter enabling imaging cells deeper (60 μm, ~10 cell diameters) in the organoid. Following background subtraction and spatial rescaling to fit training data image resolution (**Methods**), cells could be tracked through ~40 hour complex lineages without curation (**supplementary figure 3.17A,B**), even at 50 μm depth where SNR was low (**supplementary video 5 & 6)**.

We next tested performance of error prediction. We manually curated parts of both new data sets, to establish if the fractions of correct links matched the predicted link likelihoods. While for organoids on palbociclib predicted error rates were indeed well-calibrated (**supplementary figure 3.16D**), for organoids imaged on a different microscope, predictions were slightly overconfident, deviating from perfect calibration **(supplementary figure 3.17C).** However, we could recover perfect calibration simply by recalibrating the scaling temperature T used during marginalization, without retraining any of the neural networks. This recalibrated scaling temperature could be obtained by manually reviewing ~200 links (<2 hours of work, **Methods**) and did not differ between different organoids and time points **(supplementary figure 3.17D,E,F).** Finally, we note that even without recalibration, deviations in predicted likelihoods were small, with only 0.06% of links erroneously not flagged for manual curation (**supplementary figure 3.17C**). Indeed, automated lineage

analysis by error filtering gave almost identical results for cell cycle and sister analysis before and after recalibration **(supplementary figure 3.17G, H)**, suggesting that, while recalibration is generally desirable, the impact of bypassing this step is limited.

Finally, we examined performance on a non-organoid model system, using published light-sheet microscopy data of mouse blastocysts[110]. Using the above approach, we could track most cells in individual blastocysts through ~25 hour lineages, corresponding to the 16-64 cell stage, and with low error rates **(supplementary figure 3.18A,B, supplementary video 7 & 8)**. Blastocyst cells moved more rapidly than intestinal organoid cells, with displacements often larger than the typical nucleus diameter, but were still linked correctly by our tracking approach **(supplementary figure 3.18B,C,D)**. Manual correction revealed minor deviations from perfect calibration for our error predictions, that again were readily addressed by recalibrating the temperature used during marginalization **(supplementary figure 3.18E,F)**.

Overall, these results show the versatility of our approach in cell tracking across a range of model systems and imaging modalities, without requiring retraining the underlying neural networks. This contrasts with the typical workflow for machine-learning driven 3D cell tracking, where for out-of-sample data new neural networks are (re)trained[87,91,111].

3.2.10 Out-of-sample capabilities with neural network retraining

Lastly, we examined the performance of our full pipeline on an imaging data set that required retraining of the underlying neural networks. Here, we focused on a confocal time-lapse microscopy dataset of *C. elegans* embryogenesis hosted by the *Cell Tracking Challenge*[98,112], as it represented the only available confocal microscopy dataset of a 3D embryo. We trained cell detection, division, and link prediction neural networks with only minimal changes to training procedure (see **Methods**). We found that our method here performed as well as for intestinal organoid data, generating cell tracks spanning up to 7 generations (**supplementary figure 3.19, supplementary video 9)**, even though training data was limited in comparison. Upon manual review of all <99% confidence links and (dis)appearance events, corresponding to 0.9% of total links, the resulting data exactly reproduced the overall *C. elegans* lineage structure. Importantly, we also verified that our method produced well-calibrated error rates, by comparing them with the manually corrected data. Independent verification of our automated tracking results, before any correction, by the *Cell Tracking Challenge* confirmed the quality of our predictions, ranking us first in tracking performance out of 18 competitors. Our TRA-score (a metric of tracking performance with a maximum of 1[113]) of 0.995 represents a more than two-fold improvement over the closest competitors[91,92]. Overall, this demonstrates that our approach is readily applied to other multicellular systems.

3.3 Discussion

In this study we presented a conceptual innovation in cell tracking: whereas existing algorithms typically generate tracks with minimal information on correctness, OrganoidTracker 2.0 instead estimates the confidence in its predictions. Our approach exploits neural networks to predict linking and division probabilities based on 3D microscopy data, and uses a probabilistic graph representation of the tracking problem to adjust these probability estimates based on information of surrounding cells. This also enables the computing of error probabilities for any tracking feature, which report their statistical significance, thus functioning akin to p-values. This enables both highly efficient manual curation, by only correcting a minority of low-confidence tracking steps, as well as fully automated analysis, by using only high-confidence track fragments. In this way, we quantified key lineage parameters without human supervision and at high throughput for multiple organoids. The ability to compute error probabilities will be important to improving scientific transparency and reproducibility. Our method allows researchers to report the statistical significance of the cell tracking results that they produce, and the associated scientific claims, which is of increasing relevance in science. We believe these capabilities will be important in further stimulating the adoption of cell tracking methods, and the quantification of cell dynamics within tissue systems. Our approach is readily extended to cell tracking in other contexts, such as 2D cultures or embryos. OrganoidTracker 2.0 is freely available, with extensive documentation and an user-friendly GUI[53].

Our approach uses neural networks not only for cell detection, but also to estimate the probability that cells are linked or will divide between two time points. Here, the key insight is that these link and division neural networks can be trained to predict well-calibrated likelihoods, meaning that the predicted probability of a putative link or division being correct matched the probability of it being part of the manually annotated ground truth (**figures 2,3**). Predicting likelihoods necessitated a modular design, with separate neural networks for detection, linking and identifying divisions, rather than a single network for detection and linking simultaneously[87,92]. This modularity brings further advantages. First, this enabled task-specific optimization both of network architecture and training data, for instance by up sampling the number of challenging division events when training the division network. This optimization is greatly aided by the fact that these subtasks have easily interpretable probabilities as their output, that allow their isolated evaluation. Second, each network can be swapped with other implementations[111,114-118] tailored to different model systems, as long as they provide well-calibrated probabilities. Finally, it allows extending our approach with additional neural networks, to predict probabilities of other events that impact cell tracking, such as cell death, cell extrusion or abnormal divisions[119].

Our ability to predict error probabilities represents a fundamental advance in the cell tracking field. Current state-of-the-art 3D cell tracking typically rely on heuristic rules to identify tracking errors, such as flagging unrealistically large displacements or short cell cycle

times[53,87], although more systematic measures of track quality were recently developed for 3D particle tracking[120]. Recent 3D cell tracking algorithms used neural networks for cell linking[117] and detection[111] that provided approximate information on link and division probability, but not in way that supports calculating error rates and statistical significance. For 2D cell tracking, studies used approaches such as linear regression, Bayesian analysis, random forests or Kalman filters[114-116,121-123] to predict link and division likelihoods, sometimes even explicitly calibrating these outputs[115], but do not provide error rates or otherwise quantify statistical significance based on these. The key enabling step here is our marginalization procedure (**figure 3.3**), which incorporates the intuitive understanding that probabilities must be adjusted based on tracking information from surrounding cells. For instance, a low-probability link can still be of high confidence if all possible other linking possibilities are ruled out by highly probable links to other cells. Without marginalization, link probability estimates are typically under-confident, resulting in too many links erroneously ranked as low-confidence for the error probabilities to be useful for subsequent analysis (**figure 3.3**). Our marginalization procedure is independent of how link probabilities are calculated and hence can be readily incorporated in other (cell) tracking algorithms.

The current best-practice approach to address inevitable cell tracking errors is to follow automated tracking by labor-intensive manual review[87,124]. Importantly, our accurate error rate prediction strongly reduced the time required for manual correction, by focusing review exclusively on the few links with >1% error rate. Consequently, a 60 hour time-lapse movie of intestinal organoids with ~300 cells, required only 4 hours of manual review (**figure 3.4**), while this would take several days otherwise. Our predicted error probabilities also enable fundamentally novel analysis and experiments, without any human supervision: by focusing exclusively on high-confidence cell tracks or lineages, one can fully automatically extract lineage features and relationships. Using this approach, we extracted key features of cell proliferation control, such as cell cycle length, cell cycle arrest rate and cell cycle correlations between sister cells, in a fully automated manner and at high throughput (20 organoids, with a computation time of ~1 hour per organoid on a desktop computer). Automated analysis of high-confidence cell track fragments has many other possible uses. First, it can be extended to other biological events, such as cell death or cell cycle stages, when combined with fluorescent markers[109,125] or neural networks that can detect it[119]. Moreover, it enables systematic characterization of cell proliferation parameters or other features under different conditions[126], such as the addition of signaling inhibitors or drugs. Finally, it can also be used to study processes such as cell movement and rearrangement, by quantifying for example the rate at which adjacent cells separate over time[88].

Our results raise fundamental issues regarding the reporting of cell tracking-based results. For small data sets, manual curation may be performed at least on a limited number of key features such as divisions. However, for larger data sets this approach is no longer feasible, such as for larger embryo or gastruloid systems[44], or when comparing organoids across many

different conditions or genetic backgrounds. Yet, once established, reported tracking results are often treated as a given – without insight into the uncertainties. Currently the only way for readers to assess the confidence of these results and associated claims is to study the original microscopy images and perform (manual) tracking oneself, which is typically not feasible. The ability to calculate error probabilities equivalent to p-values, as we advance here, will be of general importance to mitigate this issue. Like any other form of quantification in science, such error probabilities or error probability cut-offs should be reported for displayed cell tracks, lineage trees, and for lineage features, such as cell cycles. Reporting error probabilities of published tracking data will also be crucial for data sharing, by enabling external users to assess confidence in different features of the data, even without access to the underlying microscopy images. Our work here now provides the conceptual framework and computational tools to extend this approach to a broad range of cell tracking applications.

Many salient features of aberrant organ development and homeostasis, including cancer, are only, or most readily, visible on the level of single-cell dynamics. Examples include (local) cell proliferation and death rates, spatial heterogeneity in proliferative potential, and how these are impacted by drugs and mutations. While recent studies demonstrate microscopy-based screens of cancer organoid shape and size[127], single-cell analysis at scale is not yet feasible[128,129]. Our demonstration of fully automated cell-tracking analysis is therefore of direct relevance to cancer research and opens a new frontier in understanding tissue homeostasis in disease.

3.4 Methods

Organoid culture. Mouse intestinal organoids with a H2B-mCherry reporter were used, gifted by Norman Sachs and Joep Beumer (Group of Hans Clevers, Hubrecht Institute). Organoids were grown embedded in membrane extract (BME, Trevingen) in medium consisting of murine recombinant epidermal growth factor (EGF 50ng/ml, Life Technologies), murine recombinant Noggin (100ng/ml, Peprotech), human recombinant R-spondin 1 (500ng/ml, Peprotech), n-Acetylcysteine (1mM, Sigma-Aldrich), N2 supplement (1x, Life Technologies) and B27 supplement (1x, Life Technologies), Glutamax (2mM, Life Technologies), HEPES (10mM, Life Technologies), Penicilin/Streptomycin (100U/ml 100µg/ml, Life Technologies) in Advanced DMEM/F-12 (Life Technologies). Organoids were kept in incubators at 37°C and with 5% CO2. The medium was changed every two days. Each week organoids were mechanically broken, and the fragments reseeded.

Sample preparation. Organoids were seeded around two days before imaging in 4 well chambered cover glass (#1.5 high performance cover glass) from Cellvis. In order for the organoids to locate within the lens working distance and minimize the required laser power, we placed the sample on a cold block (~4°C) for ten minutes after seeding. In this way the

organoid fragments could sink to the bottom before the gel solidified. Afterwards the BME gel was allowed to solidify at 37°C for 20 minutes before adding medium.

Microscopy. Imaging was performed on a Nikon A1R MP microscope with a 40x oil immersion objective (NA = 1.30). Around 30 z-slices with a 2 µm step size were taken per organoid every 12 minutes, with a pixel size of 0.32 μm^2.

For the low signal-to-noise data imaging was performed with a Leica TCS SP8 microscope with a 40× water immersion objective (numerical aperture, 1.10), with a pixel size of 0.4 μm^2.

Computational resources. All analysis described was done on a desktop computer with dedicated graphics card (Nvidia RTX 2080 Ti).

Intestinal organoid training data. Our training data consisted of 9 different tracked crypts together with nearby villus regions. Timelapses were between 16 hours and 65 hours long and with the full dataset totaling 281 hours (1405 frames). For a given frame around 150 cells were annotated, meaning that on the order of 200,000 cell detections and links between are present in the training data This is the same dataset used to train the original OrganoidTracker[53] so we can confidently say that any improvements are due to the new algorithm and not because of an expanded training dataset. All training data was generated in the context of, and partially used in an earlier publication[8].

General neural network training and prediction procedure. The input during both training and predicting for all neural networks consists of a list, in which each item references an image frame together with any data needed to create the final neural network input (i.e. a list of cell centers around which to crop). Only during training and prediction image frames are loaded and is the input data generated in order to minimize memory footprint. All data augmentation during training is done at run-time for the same reason. Image frames can be loaded from .tiff files but also from common platform specific file formats like *.lif* (Leica) or *.nd2* (Nikon) to avoid the need for data conversion.

Before training the neural network, the input list is randomized and split into a training and validation set (80% vs. 20%). After training the link and division detection data we perform a simple Platt scaling based on the validation dataset to make sure our predictions are well calibrated[102]. During Platt scaling we try to maximize the likelihood of the ground truth data given our scaled predictions (p^*):

$$L(p^*|x) = P(x|p^*) = \prod_i P(x_i|p_i^*)$$

With for a given link or division prediction i:

$$L(p_i^*|x_i = 0) = c(x_i = 0|p_i^*) = 1 - p_i^*, \quad and$$
$$L(p_i^*|x_i = 1) = P(x_i = 1|p_i^*) = p_i^*$$

Where the scaled predictions are given in terms of the original predictions, p, by (with A and B to be optimized):

$$p^* = \frac{1}{1 + \exp(-A \ln\left(\frac{p}{1-p}\right) + B)}$$

The maximum likelihood is then found by minimizing the cross-entropy loss between x and p^*:

$$\min_{A,B} \sum x_i \log(p_i^*) + (1 - x_i) \log(1 - p_i^*)$$

Gradient descent is done using the ADAM-optimizer for all neural networks. The full network architectures can be found on our GitHub (https://github.com/jvzonlab/OrganoidTracker).

Cell center detection – generating training data. To detect cell centers we use both the frame at the time point of interest and the subsequent frame to give the neural network access to dynamic information. We crop the images to a box which contains all annotated cell centers to avoid learning on unannotated regions. Images are then normalized after which random crops (32 x 96 x 96 x 2t) are made. Users can set arbitrary time windows and crop sizes when training their own neural networks.

To augment the data these crops are randomly flipped along the x or y-axis (50% of cases) or randomly rotated and scaled (by a random factor between 0.8 and 1.2). Further augmentation is done by randomly changing the contrast by exponentiation of the intensity values by a random number (between 0.8 and 1.2). The fluorescence intensity decay with increasing image depth can vary greatly between imaging settings. We therefore also augment the data by increasing the decay in intensity with depth by a random factor, such that the deepest frame can have up to 4-fold reduction in intensity

Cell center detection – distance map and weights. The neural network is trained to predict for every pixel in the image the distance to the nearest cell center. The distances are transformed by a Gaussian function, to give rise to diffuse spots centered around cell centers. This approach has achieved success in many cell localization algorithms where the full segmentation of cells is not available[53,92,130]. We improve this approach by also taking into account distances to nearby cells other than the closest one. By increasing the distances (and thereby decreasing intensities in the distance map) for pixels that are close to another cell, we ensure that the Gaussian spots remain well separated. The mathematical description of the 'adaptive' distance d is given by:

$$d = \frac{d_{closest}}{d_{max}} + \sum_{i \neq closest} 1 - \frac{\min(d_{max}, d_i)}{d_{max}}$$

In which d_{max} represents the maximum radius within cell centers are still relevant in computing the distance value for a pixel. It can be chosen up to the minimum distance between two cell centers before spots will overlap. The first term measures the distance to the closest cell center, while the second term increases this value if other cell centers are also within d_{max}.

The intensity values in the distance map are then given by:

$$I = e^{-d^2/2r^2}$$

In which we choose r to be $\sqrt{1/8}\, d_{max}$ to produce well separated spots.

The calculation of the distance map is done at run-time on the GPU for maximum efficiency. It can be implemented using only convolutional operations by replacing the minimum operator in the equation above by a pseudo-minimum (soft-min function).

Our algorithm allows users to only partially annotate datasets reflecting the fact that most existing manually tracked data is often focused on a limited region of interest due to time considerations. The training on partial annotations enabled by assigning large weights to pixels in the annotated regions versus the background during training. To assign these weights we change our distance map so that pixels with multiple cells nearby have lower distance values associated with them:

$$d = \frac{d_{closest}}{d_{max}} - \sum_{i \neq closest} 1 - \frac{\min(d_{max}, d_i)}{d_{max}}$$

We then use these distances to calculate the weight values:

$$W = e^{-d^2/2r^2} + b$$

where b is a small weight assigned to background pixels. By giving some weight to the background the neural network can learn to ignore debris and imaging artefacts outside of the foreground. For the intestinal organoid data b is chosen such that half of the total summed weights is associated with annotated nuclei and half with the much larger background region.

Cell center detection – neural network. The neural network used for cell detection is very similar to the 3D-UNET used in the previous OrganoidTracker[53]. The different time-points in the input are treated as different channels. A new element in the network is a final smoothing layer (convolution with Gaussian kernel with a 1.5 pixel width). Because the center point annotation is inherently noisy (not pixel perfect) the predicted output should be smooth. By enforcing this explicitly we reduce overfitting and speed up the training.

Cell center detection – peak finding. From the predicted distance map, we localize the cell centers by using a peak finding algorithm, as described before[53]. Peaks within a certain radius (half the typical distance between nuclei) of other higher peaks are excluded by the peak finding process to avoid oversegmentation due to noise in the predicted distance map.

During cell division cells round up and their distance to other nuclei increases. At the same time cells are more prone to oversegmentation as the H2B fluorescence is not uniformly distributed anymore because of chromosome condensation. To counteract this, we revisit the cell detections after we have predicted the division probabilities (see below) and merge dividing cell detections (defined as having a division probability greater than 50%) that are closer than 5 µm from each other.

Cell center detection – Evaluation. Cell center detection was evaluated as previously done[53]. We compared predicted data with partially annotated manual datasets. The evaluation data consisted of five different organoids, imaged on different days, for which at least one crypt was fully tracked. The organoids were tracked for between 90 and 320 frames.

For every cell center in the manually annotated dataset we check if there is a predicted cell center within 5 µm, these count as true positives. A predicted center can only match a single cell center in the manual data. Unmatched manual annotations are false negatives. Predicted cell centers that remain unmatched and are within the manually annotated region (5 µm distance from an annotation) are counted as false positives. Consigning the evaluation to annotated regions means that mistakes far from the epithelial layer are ignored (i.e. debris recognized as nucleus), but these are both rare and generally irrelevant for tracking.

Recall is calculated by dividing true positives by the total number of manual annotations. Precision is defined by dividing the false negatives over the amount of predicted cell centers within the annotated region. Accuracy is the number of mistakes over the sum of all observations (true positives, false positives, false negatives).

To test the effect of our 'adaptive' distance map we also trained a network on a target mapping that consisted simply of Gaussian spots around the cell centers. In order for these spots to not overlap we had to half their radius relatively to the 'adaptive' version. The pixel weights were kept the same (**supplementary figure 3.2**).

Cell center detection – Cellpose comparison. We used the CellPose 3D module[118] to produce nuclear masks for three timeframes of our test dataset. The CellPose algorithm was run from a dedicated OrganoidTracker plugin. We used an expected nucleus diameter of 25 pixels. After obtaining nuclear masks, we computed the centroid positions as the center-of-mass of each 3D mask. We manually removed CellPose centroids that correspond to over- or undersegmentation. These validated centroids were then compared to the OrganoidTracker predictions. The analysis was limited to a tissue depth of 15 µm, with poor CellPose segmentation for higher depths.

Link detection – proposing possible links. To avoid looking at extremely implausible links we propose links based on the distance between the subsequent cell detections. During both training and prediction, we only consider links from a cell detection to a cell detection in the next frame that are at most two times further away in distance than the closest cell in the next frame.

Link detection – generating training data. The input of the neural network for link prediction consists of a crop centered around a cell center, a crop around the cell detection in the subsequent frame and a vector describing the distance in pixels between the cells. The two crops are 16x64x64 in size and both contain the two time points containing the cell center detections. Users can set arbitrary time windows and crop sizes when training their own neural networks.

Data is augmented in the same way as during cell center detection, except that we do not vary the decay in intensity with depth as the crops are much smaller in the z-dimension. Instead we increase the range in which we vary contrast (exponentiation by a number between 0.5 and 1.5).

To aid prediction we provide the neural network with direct information about the direction of movement by adding the displacement vector to the neural network inputs, besides the crops around the cell centers. It is known that convolutional neural networks have trouble integrating information in the form of Cartesian coordinates[131]. We therefore add an extra three channels to both crops. These contain for each pixel the x, y and z distance respectively to the other cell center detection in the proposed link.

We upsample difficult cases, cells that are dividing (within a window of an hour around cell division) or move a considerable distance (more than 3 μm less than 7), by replicating these 5 times in our training data.

Link detection – neural network. The first part of the neural network for link detection consists of two convolutional neural networks (CNNs). To maximize the amount of information extracted one CNN takes in the concatenated crops while the second CNN takes as input a single crop (two identical copies of the second CNN are available to analyze both crops). This means that one CNN can integrate pixel information between crops and directly assess how alike the two cell detections at subsequent time points are. The other CNN is forced to focus on a single crop, which could in combination with information about the direction of movement already be enough to assess the link probability.

The features extracted by the CNNs in combination with the displacement vector are then fed into multiple densely connected neural network layers to yield a prediction.

Link detection – Evaluation. To evaluate the link neural network, we used the same set of evaluation data as used in evaluating the cell center detection. See main text for the evaluation procedure.

To test the effect of adapting the training data we also trained a link detection neural network without upsampling difficult cases (see ‘Link detection – generating training data’). We then compared accuracy, precision and recall across all evaluation organoids (**supplementary figure 3.4**).

Division detection – generating training data. The input of the division detection neural network is a crop (12x64x64) centered around a cell center, with the previous and subsequent

frames included for dynamic information. Data augmentation is done the same as during link detection training.

To avoid a too low frequency of images related to cell division, we upsample cells in the process of division (within a one hour window around the nucleus dividing) by replicating them 10 times in our training data. We also upsample all dying cells (cells with tracks ending before the end of the experiment), as these can closely resemble dividing cells. From all other cell detections, which are often trivial to predict as non-dividing, only a random subset is included so that they make up 20% of the total dataset.

Division detection – neural network. The design of the division detection neural network mimics that of the link detection network. A convolutional neural network extracts features that are then fed into a dense layer to generate the prediction. The main difference is that due to limited nature of the division datasets (there are only hundreds of divisions present in our training data) we employ only a single dense layer to avoid overfitting.

Division detection – Evaluation. To evaluate the division neural network, we again used the same set of evaluation data used in evaluating the other neural networks. See main text for the evaluation procedure.

To test the effect of adapting the training data we also trained a division detection neural network without upsampling difficult non-dividing cases (see 'Division detection – generating training data'). We replaced these difficult cases by randomly selected cell centers, so that divisions make up the same fraction of the training data as in our normal training procedure. If we would truly train on an unbiased sampling the data, so that non-divisions make up the vast majority, this would cause the training procedure to not converge. We then compared accuracy, precision and recall across all evaluation organoids (**supplementary figure 3.5**).

Graph description. In our graph description of the dataset we follow the framework developed by Haubold et al.[99]. Here the nodes of the graph are the detected cell centers and the edges the proposed links. These edges have an associated energy penalty that is the relative negative log-likelihood that the link is true as predicted by the neural network. The nodes have an associated division penalty which is again the negative relative predicted log-likelihood.

Within this framework we also have to assign energy penalties to the events where a track disappears or appears or that a cell detection is a false positive. A track can disappear when a cell dies or its next position is not detected. The disappearance probability is thus the combination of the death rate and the false negative rate of the neural network. Here, the latter makes the dominant contribution. Tracks can appear when their previous position is not detected, which again relates to the false negative rate. The probability of a cell detection being spurious is given by the false positive rates. All these rates can be estimated from the validation of the cell detection neural network and are around 1%. Varying these probabilities

within an order of magnitude (3% to 0.3%) does not significantly affect the track prediction or the marginalization procedure (not shown).

To account for cells appearing or disappearing because they are close to edge of the imaging volume and can leave the imaging volume, we assign lower (dis)appearance penalties (corresponding to a 10% chance of (dis)appearance) to cell detections at the edges of the volume.

One could imagine a neural network that would assign explicit probabilities to the correctness of cell detections so that we could use node specific (dis)appearance penalties. This should lead to minor improvements in track quality, but such an approach would have several drawbacks. First of all, training data is limited because the cell detection network makes few mistakes. Furthermore, such a neural network would have to be retrained every time a new cell detection network is trained as it is specific to the type of mistakes that that network makes. Integrating a neural network to identify dying cells and adapt the disappearance probabilities accordingly would be more feasible[119], but of limited use due to the rare nature of cell death in our system.

In principle the predictions made by division and link detection neural networks are probabilities conditional on the correctness of the underlying cell detections, because only correct cell detections are in the training data. It is possible to assign energy penalties in such a way that they represent probabilities of a link or division conditional on the existence of the node it is coming from, by combining the chance that a link is incorrect and that its source node does not exist in a single energy penalty. This could avoid including some oversegmentations that persist over multiple subsequent frames and have high probability links between them in the tracking solution. But including correct links between oversegmented cells is in our case actually the preferred behavior. Not including these links would hamper our approach of solving these oversegmentations during post-processing (see below). This does mean that after marginalization we also have to interpret the predicted error rates as the chance that the two different cells associated with the detections are not linked, not the chance that the link is ‘incorrect’ because one of the two detection is due to an oversegmentation. Because oversegmentations on its own already introduce errors per definition, as a track caused by oversegmentation both has to appear out of nowhere and disappear again, this will not cause any missed errors.

Flow solver. We use the flow solver developed by Haubold et al.[99] to find the most likely set of tracks. In order to help it converge to an optimal solution we prune the graph of high energy edges. We do this by comparing every edge to its alternatives; links having the same source or target nodes. If a link with a much lower penalty is available (>4.0 difference, corresponding with a 10,000 times more likely link) we remove the edge. This was not done during the marginalization evaluation (**figure 3.3**), where link removal like this would introduce a bias in the non-marginalized probabilities for very unlikely links. Potential divisions that have a probability below 0.01 are also removed.

The flow-solver has sometimes trouble converging or halts prematurely especially in the presence of a large number of low certainty predictions. To circumvent this, users can also use the Viterbi-style algorithm proposed by Magnusson et al.[132] as implemented by Haubold et al.

Fine-tuning flow solver solution. Because the flow solver does not guarantee an optimal solution, we fine-tune our solution by checking for every link if removing it and replacing it with an appearance and a disappearance would lower the total energy. We then also look at pairs of links in the solution that connect two nodes at timepoint *t* with two nodes at timepoint *t+1* and check if they should be replaced with a pair of edges that connects the nodes the other way around. We perform 3 cycles of this pruning and swapping of links.

Solving over- and undersegmentation. Our probabilistic description allows us to add and merge nodes in the graph in a statistically rigorous manner to tackle the track fragmentation caused by over- and undersegmentation. The procedure relies on four key parameters: the cell detection network's false positive and false negative rate, and each cell's predicted link and division probabilities. The false positive and negative rates follow directly from the performance of the cell detection neural network, while the link and division probabilities are predictions from the linking neural network and the division neural network. Hence, these parameters are in principle obtained through the network training procedure, without any further user intervention.

Oversegmentation occurs when a single cell generates two or more cell detections, potentially during multiple frames, causing tracks to split up erroneously. Such split tracks are identified as follows: these pairs of tracks should partially overlap in time (minimum of one and maximum of three frames) and nodes in the different tracks should be connected by relatively high-probability edges, that are otherwise not part of the tracking solution. This reflects that, if the tracks represent the same cell, edges between nodes in the two tracks should be likely. If the combined probability of an edge connecting the two tracks and the probability of a false positive cell detection, as given by the false positive error rate is higher than the probability of a track disappearing and another appearing (based on the false negative rates), we connect the tracks and prune the overlapping cell detections (**supplementary figure 3.6A**).

Undersegmentation occurs when a cell is not detected, leading to a single track becoming fragmented into two tracks. We identify fragmented tracks with a single frame gap between them, and propose a new node that connects the tracks only if their start- and endpoints are within a sufficiently short distance. Here, a cell detection is considered near to another one, if it is one of the six closest neighbors. The added node receives a 3D position that is the average of the positions of the start- and endpoints of the two tracks, and is assigned a probability of being correct that is equal to the false negative error rate. In the graph containing all potential links, new edges are then made to all nearby nodes, with an energy penalty representing a uniform link probability (**supplementary figure 3.6B**).

On a practical level, the post-processing is implemented by first identifying all situations where cell tracks appear or disappear. The algorithm then first addresses oversegmentations, by attempting to connect appearing tracks with a nearby disappearing track that overlap in time for a maximum of three frames. After that, the algorithm addresses undersegmentation, by attempting to connect appearing tracks with a nearby disappearing track that has disappeared just one timeframe before.

Fundamentally, our post-processing solves a fundamental drawback in graph-based tracking frameworks which treat every cell detection as independent evidence for the existence of a cell. If, for instance, a cell is oversegmented in multiple subsequent timepoints this is treated as very strong evidence that there are actually two cells present. It is obvious that this actually confers little more evidence than a single oversegmentation because these detections are and should be highly correlated between frames. Revisiting potential oversegmentations during post-processing allows us to treat multiple subsequent oversegmentations as a single false positive event. Our undersegmentation correction method solves another problem with using flow solvers for tracking; they can ignore cell detections when making tracks but cannot add nodes for missed cell detections. A priori it is difficult to determine where 'helper' nodes might need to be added and allowing cell 'merging' to deal with undersegmentation[99] makes the tracking problem much less constrained. We instead solve it with an easily understandable and straightforward post-processing step. Earlier cell tracking solutions have employed conceptually similar methods but have to rely on manually picked parameters to regulate post-processing in the absence of a probabilistic description[104]. In contrast, we use our probabilistic graph description to rigorously identify the proper post-processing steps with minimal need for user set parameters.

There is in principle no need for the user to adapt the post-processing procedure for different datasets, as long as the neural network predicted probabilities are well-calibrated and the user-set (dis)appearance probabilities are realistic (see the 'graph description' section). Similarly, retraining the neural networks for a new dataset automatically ensures proper post-processing on the new dataset as well. When reusing already trained neural networks in a new context, it can be necessary to change the (dis)appearance probabilities to reflect the performance of the cell detection neural network in this context.

Marginalization. Marginalization is done on a subset of the graph to make it computationally tractable. We assume that the most informative edges (and their associated nodes) are between the same time points as the link of interest and are the ones closest to it in space. Distance is measured by how many steps on the graph have to be made to traverse edge-wise from the target node of the link of interest (**supplementary figure 3.7B**). Taking three steps as cut-off for inclusion in the subset yielded a computation time for marginalization similar to that needed for neural network prediction of link and division probabilities, ~1 hour for an imaging experiment of over 300 frames with over a hundred cell detections per frame.

The number of steps used to construct the subgraph can be changed by the user. We find that going beyond three steps, which already includes all the links of neighboring cells, does not meaningfully improve prediction quality (**supplementary figure 3.7C**). This lack of improvement can partly be explained by marginalized link predictions for which the subgraph does not change when increasing from 3 to 4 steps (~30% of cases). For ~20%, the subgraph simply has no connections beyond 3 steps. For the remaining ~10%, the 4 step subgraph has too many elements to be evaluated in a reasonable time and we are forced to use the 3 step subgraph instead. Here, we use a cut-off so that no more than $\sim 2^{16}$, possible tracking solutions have to be checked. However, for links where the subgraph used does grow, we still see little change in the prediction. This is mostly because many of the predictions with the 3-step subgraph were already very high-confidence (64% of cases are above 99.99% or below 0.001%), suggesting most contextual information was already incorporated. Any further improvement thus made little difference for the prediction quality as measured by the cross-entropy loss.

For every node in the subset all edges that point to non-members of the subset are combined in a single edge that accounts for the total probability to connect to a node outside of the set.

After subset selection we construct a set of potential tracking solutions, test which solutions fit the graph constraints and calculate their associated energy. To avoid having to check the full set of binary combinations of events ($\sim 2^N$) we construct the set by varying for every target node in the *t+1* timepoint which node in the previous timepoint *t* is connected and combining all these variants. In this way the number of constructed potential solutions scales as $\sim (N_L/N_T + 1)^{N_T} + 2^{N_S}$, in which N_T is the number of target nodes, N_s the number of source nodes, and N_L the number of edges. The target nodes can on average contact N_L/N_T possible source nodes and can appear without a source, while the cells represented by the source nodes can either disappear or not. We will refer to these possible variants as 'microstates' of which we will later combine the probabilities in order to compute the error rates.

Microstates can be encoded as a vector with as its length the number of events (1 if an event, like a link or division, is part of it, 0 if not). To check if a microstate is possible, we can construct a matrix that encodes the flow constraints on the graph. This matrix gives the net flow into every node when multiplied with a microstate vector. An outgoing link or disappearance event represents a flow of -1 while an incoming link or appearance gives a flow of 1. Divisions are represented with a -1 flow as they should allow an extra outgoing link. When for one or more of the nodes the flow is unbalanced, the microstate is rejected and excluded from the partition function. Total energies are calculated by taking the inner product with a vector containing the energy penalty per event. These energies are then divided by the 'temperature' for proper calibration (see next section).

The probability of a link of interest (A) being true given all predictions made on the elements of the subgraph (G) is thus found by normalizing the probabilities associated with microstates

containing that link by the sum of the probabilities of all possible microstates. The probability of a given microstate (W_A) in turn is proportional to exponent of the negative sum of the energy of all its elements ($E(W_A)$).:

$$p(A|G) = \frac{\sum_{W_A} e^{-E(W_A)/T}}{\sum_{W} e^{-E(W)/T}}$$

To reduce the computational burden, links that are deemed almost certainly correct (>99.99%) or incorrect (<0.01%) are marginalized over a minimal subgraph containing only the other input edges of the target node of the link in question. When the estimated number of microstates that would need to be constructed exceeds 2^{16}, we shrink the subgraph by one 'step' to avoid long computation times.

Motivation for using 'temperature scaling'. The marginalization procedure without temperature scaling assumes that the energy penalties are derived from information that is unique to the predictor, a neural network in our case. This is not a realistic assumption as predictions might be made on the basis of overlapping crops and on shared baseline estimates. Not accounting for this overlapping information leads to overconfidence (**supplemental figure 3.7**).

In our solution for this problem, we propose to split all predictions in a component that is based on information shared between neural networks and in one based on information unique to that prediction. The predictions (p) can then be seen as the product of the relative probabilities based on this shared and unique information:

$$\frac{p}{1-p} = \frac{p_{shared}}{1-p_{shared}} * \frac{p_{unique}}{1-p_{unique}}$$

This allows us in turn to split up the energy (the negative relative log-likelihood) in a shared and unique component. We then assume that the energy related to the probability based on the shared information estimate is proportional to the total energy (E_i). This assumption reflects our intuition that a neural network prediction should be directionally reflected by both the unique and shared component. If for instance a link is highly likely, then this can probably be deduced both from the shared and the unique information available to the network and both energies should be highly negative. This gives:

$$E_{i,unique} = E_i - E_{i,shared} = E_i - aE_i$$

With a a constant between zero and one.

When calculating the energy of a microstate (E_W^*) we can then sum the unique energies while assuming we can combine the shared information in a weighted manner. This weighing factor b (smaller then 1) should be low if all the shared information is shared between all events

and higher if the overlap is less (for instance when a prediction made about a link mostly shares information with adjacent links, but not with all elements in the subset):

$$E_W^* = \sum_{i\ in\ W} E_{i,unique} + bE_{i,shared} = \sum_{i\ in\ W} (1-a)E_i + baE_i$$

From this we derive that we can account for shared information by using a single factor that functions as a temperature (T). This temperature is high if much of the information in any given prediction is not unique (high a) and if this shared information is shared with all other predictions (low b):

$$E_W^* = (1-a-ba) \sum_{i\ in\ W} E_i = \frac{1}{T} \sum_{i\ in\ W} E_i$$

$$with;\ T = \frac{1}{1-a+ba}$$

Marginalization as an opinion pooling procedure. We can also motivate our marginalization procedure without relying on analogies with statistical physics. Instead, we can interpret our method as an extension of the 'multiplicative opinion pooling' framework as proposed by Dietrich[105,106]. The idea of combining predictions in a machine learning context has an older history[133], but the specific framework of Dietrich and List enables us to neatly deal with prior probabilities and overlapping information. This will prove to be key in producing well-calibrated outputs.

Multiplicative opinion pooling suggests that opinions of different agents (different predictions by neural networks in our case) can be combined by multiplying them:

$$P(\omega) \propto \prod_i P_i(\omega)$$

In which ω is a state in the set of possible states Ω and P_i denotes the probabilities predicted by individual predictors. Shared information between predictors can be incorporated in this framework by normalizing the predictions by the predictors' priors based on the shared information. Conceptually this means that predictors first arrive at a consensus P_0 on the basis of their shared prior information after which their unique information is pooled multiplicatively.

$$P(\omega) = c\ P_0(\omega) \prod_i P_i(\omega)/P_{i,prior}(\omega)$$

With c functioning as a normalization factor:

$$c = \frac{1}{\sum_\omega P_0(\omega) \prod_i \frac{P_i(\omega)}{P_{i,prior}(\omega)}}$$

In this framework each predictor must have an opinion on all possible states. In our case predictors make only a single prediction on an event a (a link or division) that is part of a state. Therefore, we redefine multiplicative pooling as:

$$P(\omega) = c\, P_0(\omega) \prod\nolimits_{a \in \omega} p_a / p_{a,prior} \prod\nolimits_{a \notin \omega} (1 - p_a)/(1 - p_{a,prior})$$

In which the microstate probability is now proportional to the product of the probabilities that its constitutive parts are true and the other events are false. The probability of a given event can then simply be calculated as:

$$P(a) = \sum\nolimits_{\omega \ni a} P(\omega)$$

By extending multiplicative pooling in this way we retain a major motivation behind multiplicative pooling, namely 'individualwise bayesianity'. This axiom states that is should not matter to the final prediction if extra information is integrated before or after the pooling procedure, as the input information is the same. In our case this holds on two levels (see supplementary text for the proof). Firstly, it does not matter when we introduce information about a microstate when calculating its probability ($P(\omega)$). It also does not matter when information about an individual event is introduced when we are calculating its probability ($P(a)$). This provides large flexibility in post-hoc integration of new opinions, such as the judgement of a human reviewer.

The question remains how to extend our concept of 'temperature' to this framework. For simplification we can rewrite everything in terms of relative probabilities ($L_i = p_i/(1 - p_i)$):

$$P(a) = c \sum\nolimits_{\omega \ni a} L_{0,\omega} \prod\nolimits_{i \in \omega} \frac{L_i}{L_{i,prior}}$$

$$with;\ c = \frac{1}{\sum_{\omega}\ L_{0,\omega} \prod_{i \in \omega} \frac{L_i}{L_{i,prior}}}$$

The question now remains how to define the consensus prior L_0 and determine the priors. Dietrich and Listz suggest using a geometric mean on the priors if the shared information is completely shared between all agents[106]. In our case this is not necessarily true so we let the weight associated to a single prediction be free (b) instead of $1/n$. For the priors we again assume that the shared information is proportional (with a factor a) to the total information held by an agent.

$$P(a) = c \sum\nolimits_{\omega \ni a} \prod\nolimits_{i \in \omega} {L_{i,prior}}^{b} \prod\nolimits_{i \in \omega} \frac{L_i}{L_{i,prior}}$$

$$P(a) = c \sum\nolimits_{\omega \ni a} \prod\nolimits_{i \in \omega} {L_i}^{ab} \prod\nolimits_{i \in \omega} {L_i}^{1-a}$$

$$P(a) = c \sum\nolimits_{\omega \ni a} \prod\nolimits_{i \in \omega} {L_i}^{1/T}$$

$$with;\ T = \frac{1}{1 - a + ba}$$

Which is equivalent to the description we arrived at using the statistical physics framework.

We finally want to contrast this opinion pooling procedure with updating a 'Bayesian belief matrix', a (cell) tracking approach that uses link probability estimates to connect non-dividing object detections[114,134]. This method cannot integrate division probabilities and can only take one type of constraint into account; the fact that cells cannot merge. In situations where these are the only constraints present (for instance when considering a subgraph where only one cell is present in the later timepoint) we show that this approach is equivalent to our marginalization method (see supplementary text).

Estimating the calibration temperature. We find the optimal temperature (as defined by the binary cross entropy loss) by calibrating on the training data. To do this we use the neural networks to predict link and division probabilities for the cell detections in the training data. Then we perform the marginalization and compare the marginalized link probabilities to the manual tracking. The task is now to find a 'temperature' (T), for which the predictions p_i are closest to the ground truth (l_i denotes the truth value of a link). That is the temperature for which the likelihood of the ground truth given the predictions is maximized and the binary cross entropy is thus minimized:

$$\min_T \sum l_i \log(p_i) + (1 - l_i) \log(1 - p_i)$$

With p_i given by:

$$p_{i,T} = \frac{\sum_{W_A} e^{-E(W_A)/T}}{\sum_W e^{-E(W)/T}}$$

In practice most of the energy contribution in our marginalization comes from a handful, often two microstates. As an example, the dominant microstates of an uncertain link often take the form of an option where all cells move half a cell to the left and another where they move half a cell to the right. For a given link A one of these states dominates the microstates that contain the link (W_A) and the other the ones that do not contain it $(W_{\setminus A})$. This allows us approximate the marginalized probability $p_{i,T}$ as:

$$p_{i,T} = \frac{1}{1 + \frac{\sum_{W_{\setminus A}} e^{-E(W_{\setminus A})/T}}{\sum_{W_A} e^{-E(W_A)/T}}}$$

$$p_{i,T} \sim \frac{1}{1 + \left(\frac{\sum_{W_{\setminus A}} e^{-E(W_{\setminus A})}}{\sum_{W_A} e^{-E(W_A)}} \right)^{-1/T}}$$

$$p_{i,T} \sim \frac{1}{1 + \left({}^{p_{i,T=1}}/_{1 - p_{i,T=1}} \right)^{1/T}}$$

$$p_{i,T} \sim \frac{1}{1 + {L_{i,T=1}}^{1/T}}$$

This clearly and conveniently maps on a linear regression problem, where we have to learn parameter $1/T$ given the original marginalized relative likelihoods $\left(L_{i,T=1}\right)$ as an input.

The temperature obtained this way works well on data that is part of (**supplementary figure 3.7**) and outside of (see next section) the training dataset in producing well-calibrated error rates. This proves that the simplifications made to arrive at a single correction factor amenable to linear regression are allowable.

The obtained temperature (like the Platt scaling parameters previously) is used in our algorithm as a point estimate without considering the uncertainty of associated with the calibration. This can be justified by the tight confidence intervals we obtain (**supplementary figure 3.7D**) and the robust calibration we see across datasets (**figure 3.3, supplementary figure 3.10**). For a fully Bayesian description of our framework, that includes calibration uncertainty, see the supplementary text 'Full Bayesian description of link error prediction framework'

Evaluation of marginalization procedure. To evaluate the correctness of the marginalized error rates we again compare our predictions against the five fully manually annotated organoids used for the other evaluations. We use the manually annotated cell centers as the input for our division and link detection and perform marginalization afterwards. This allows us to compare all error rate estimates to a fully human derived ground truth, without the need to map machine predicted cell centers on the cell centers annotated by humans. These mappings are not trivial and errors in these mappings can strongly skew the results. Furthermore, because the human assigned links are completely independent from the algorithm output, we deem this the strongest test for our marginalization procedure.

We bin the marginalized link predictions in groups based on their relative log-likelihood (15 bins). For every bin we then compute the average probability of the link being correct and compare this to actual amount on correct link in this bin as determined by looking at the ground truth.

We also perform this evaluation on the manually reviewed tracking data (see next section). Here we started out with cell centers predicted by a neural network. Verifying that the error rates are well-calibrated in this case shows that the marginalization procedure is not dependent on human annotated cell centers. We compare the error rates against tracking data where all links are corrected, but no under- or oversegmentations are solved (**supplementary figure 3.9**). Solving segmentation errors involves changing the graph representation and thus introduces links without an associated error rate prediction, making evaluation impossible.

Manual review. In order to evaluate manual annotation, we reviewed the possible errors for 3 complete organoids tracked for around 100 to 300 frames. Potential errors were flagged at all links that had a marginalized probability below 99% and the start and endpoint of appearing and disappearing tracks respectively. We first corrected all potential link errors and used the corrected data to check the calibration of the marginalized predictions as described above. We then checked all other errors and identified their cause for the largest (>300 frame long) dataset.

Error correction was done in our GUI[53], that zooms in on errors and tells user about the kind of error they encountered; possible link mistake, track appearing or track disappearing. The GUI also allows backtracking, i.e. the selection of cells of interest, based for instance on their cell type or final position, to focus curation only on these cells.

Palbociclib intestinal organoids tracking. Palbociclib (at final concentration of 10 μM) was added two days after seeding and organoids were then imaged for two days. Three crypts that stayed in the field-of-view for the full image duration were chosen for analysis.

Image preprocessing. For out-of-sample usage of our intestinal organoid trained neural network we have identified two key preprocessing steps to improve tracking results: scaling and background subtraction. Regarding scaling, convolutional neural networks (and UNET's by extension) are generally not scale free. So, in order to avoid over- or undersegmentation the nuclei size should match the nuclei sizes in the training data. This does often, but not necessarily have to correspond to matching the pixel resolution.

Secondly during the acquisition of the intestinal organoid data the detector gain and offset was set such that the background (meaning the region outside the organoid) largely had fluorescence values of zero. Subtracting the background so that this holds for the out-of-sample dataset as well helps to restrict the cell detections to the region containing the tissue.

Lastly, we have seen that for data where cells have large differences in nuclear fluorescence (unpublished) it helps to reduce the contrast using a gamma correction.

Recalibration. Recalibration of the error rates for out-of-sample data follows the same process as the initial estimation of the scaling temperature (see **estimating the temperature scaling** section): we compare the marginalized predictions against a ground truth dataset to find the optimal temperature that minimizes the cross-entropy loss between predictions and truth values.

In order to create the ground truth, the user has to correct potential mistakes in a number of representative frames. We find that correcting around 200 potential linking mistakes is generally enough to get tight estimates of the new scaling temperature. The procedure requires users to review all potential errors in a frame to avoid bias in which mistakes are corrected.

The recalibration procedure thus functions like this:

1. Predict tracks and compute error rates
2. In the manual curation GUI select frames for which to correct potential mistakes (aim for more than 200 potential mistakes)
3. Correct mistakes in the GUI
4. Recalibrate the error rates using the temperature scaling functionality
5. Recompute the error rates with the new scaling temperature

Low-SNR intestinal organoids tracking. The low SNR-intestinal data was taken (but not yet analyzed) in the context of Zhang et al. and imaging was done as described there[88]. Pre-processing consisted of downscaling in XY by a factor of 1.33 and background subtraction using a tophat filter. Post-processing was changed to retain deep tracks, up to 60 µm deep in the tissue.

Blastocyst tracking. The two longest timeseries in the BlastoSPIM dataset[110] were chosen for analysis. Pre-processing consisted of downscaling the image in XY by a factor of 2.5 and background subtraction by subtracting a constant value. This downscaling was not meant to match the resolution of the blastocyst data to the training data (different by a factor of 1.25 in XY), but rather to match the nuclear volumes, which are considerably larger in the blastocyst (~6.5 µm nuclear radius vs ~3.5 µm in the intestinal organoids). When a blastocyst underwent a major rotation, we ignored that timepoint in our analysis of the error rates. These major rotations occurred only during two frames in only one of the blastocysts. To correctly track cells through these major rotations our tracking algorithm would simply have to be combined with an image registration steps, as was done in the original paper describing the dataset[110].

***C. elegans* cell tracking.** We obtained the *C. elegans* embryo datasets from the cell tracking challenge website[98]. The available training data consisted of two fully annotated movies (~150 frames long) following cells from the 2 to ~128 cell stage.

We trained new cell detection and link and division prediction neural networks on the two provided annotated training datasets. As all cells in the imaging volume were annotated there was no need to crop the image during cell detection training and we could increase the background weighting to 0.95 without risking training on unannotated cell centers. Due to the difference in nucleus sizes compared to the intestinal organoid data we also increased the radius parameters in the distance mapping for the cell detection. All other networks were trained as for the intestinal organoids. We estimated the proper scaling temperature for the marginalization by calibrating on the training data as described for the intestinal organoid data.

We used one of the unannotated ‘challenge’ datasets to evaluate tracking quality and validate that the marginalized probabilities were well-calibrated. We did this by manually checking all potential errors and using the corrected dataset as our reference.

Automated lineage dynamics analysis. For the automated analysis we first filtered out all links that had a marginalized probability below 99%. All tracks that do not end in a division are considered censured.

A key assumption underlying survival analysis is that the probability of an event happening is independent of the chance of being lost to follow up. In our case this assumption is broken as cells are relatively often lost when they are close to dividing (due to rapid nucleus movement) and cell division is the key event when studying lineage dynamics. This means we would underestimate the number of dividing cells, because we tend to lose track of them just before they divide. We break this dependency by using a division detection neural network to check for every track that is lost to follow-up if it is lost during the division process. We then reassign tracks which end in a predicted division (>50% predicted probability) from the censured category to the divided class. Now observing the division events is no longer affected by uncertainties in the tracking during the division process.

The neural network trained for this task was trained in the exact same way as described before except that we are not interested in pinpointing the exact moment of chromosome separation. We want also want to classify tracks as dividing if they are lost during any other moment of the division process. We therefore classify all cells within two frames around division as dividing during training. Because of the varying length of the division process, we exclude time points directly around this window to avoid including cells in the training data that look clearly mitotic but are just outside the window.

We can also use this neural network to split tracks that contain a division, but where in initial tracking no division was assigned due to lack of a plausible daughter cell, for instance because one of the daughters moves out of view. Therefore, we break up tracks when the chance of division is on average higher than 99% for three consecutive frames.

Our method detects some cells with very short cell cycles, where the cell division generally leads to cell death and not in two daughter pairs, potentially reflecting polyploid cells. These are not classified as dividing in the manually annotated data so we remove these very short cell cycles (below 6 hours). This has the added benefit that it also removes some cases where the division neural network wrongly assigns a division to a track end. Although the chance of that happening is low; it happens generally in less than 2,5% of tracks.

For survival analysis we use the ‘surv’ package in R and for the fitting ‘survflexcure’. During analysis we only use tracks that start in a division and use the next division as the event under study. Cell cycle times are analyzed by fitting a Gaussian hazard to the data allowing for a ‘cured’ fraction, that will not divide again. The mean of the Gaussian represents the average cell cycle and its standard deviation the spread around this mean. The ‘cured’ fraction is used as an estimate of the fraction of differentiating cells. Prior to fitting we remove outliers that are more than ~7 hours from the mean (more than 3 times the standard deviation). To avoid

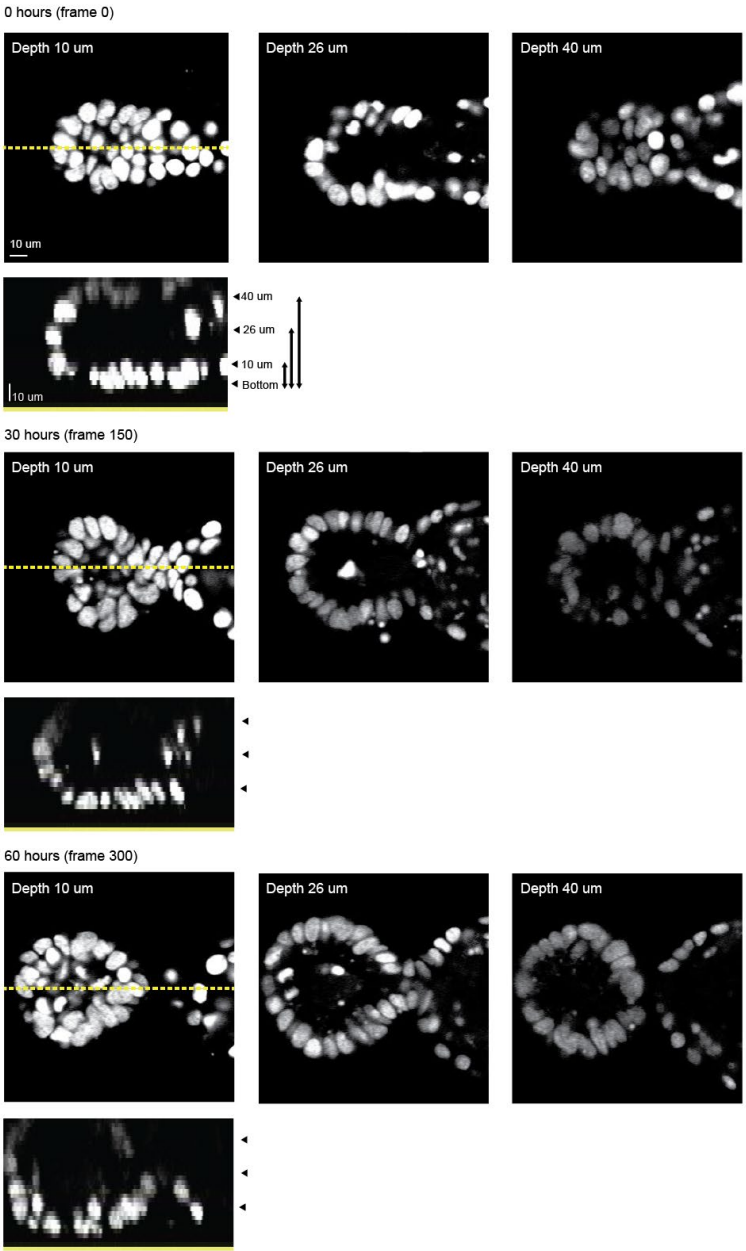
dealing with negative times we fitted a log-normal distribution to the exponents of the survival times instead of using a normal distribution directly.

Manual data for comparison is analyzed in the same way, but the only censoring events derive from cell death, the end of the experiment or cells leaving the imaging volume. No neural network thus has to be used to check if censured tracks end in a division.

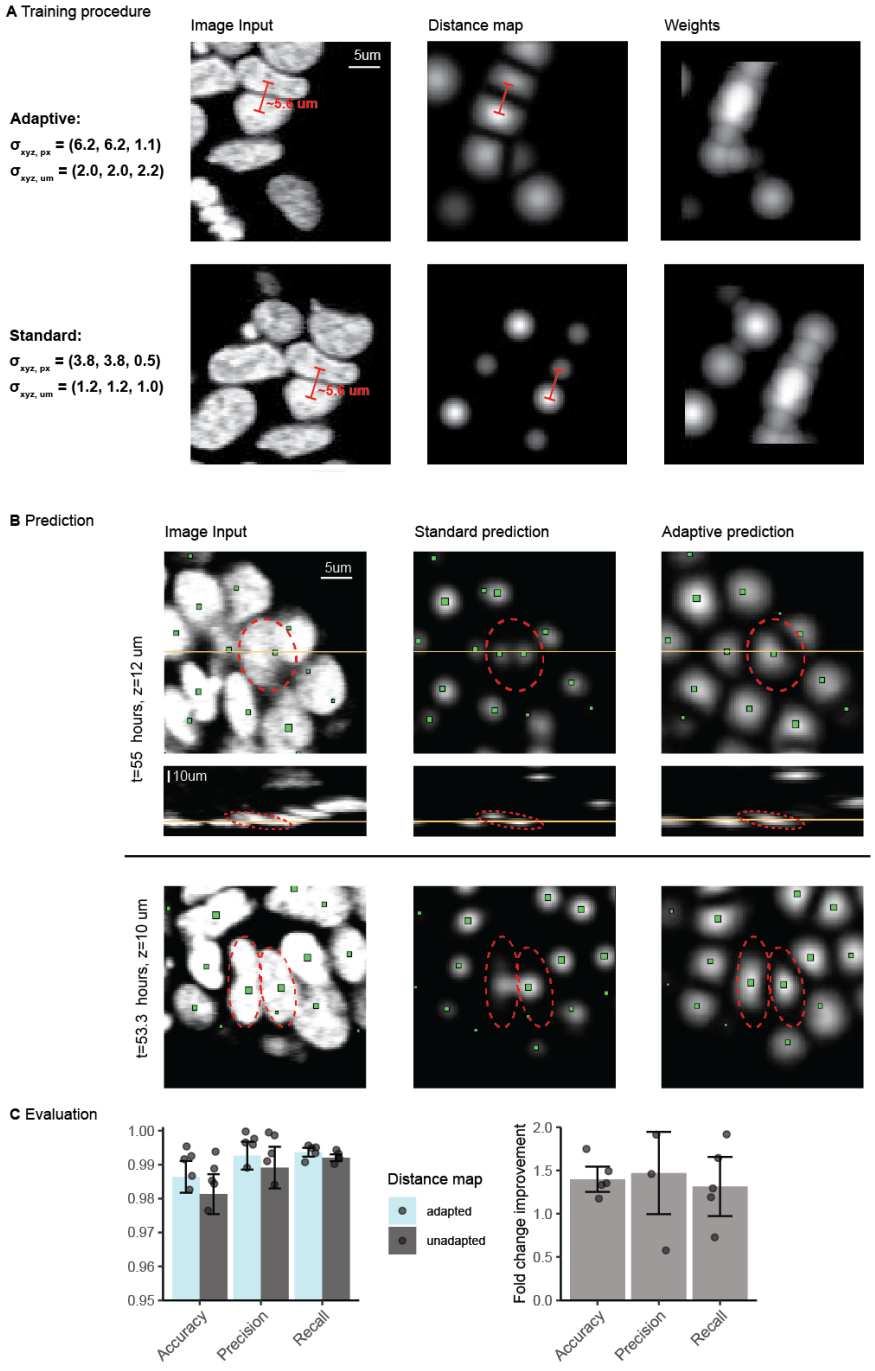
Data availability. Due to storage considerations all imaging data and accompanying tracking data are available upon request. Sample imaging and tracking data is available on Zenodo (DOI: 10.5281/zenodo.13982844).

Software availability. The OrganoidTracker software is freely downloadable from GitHub (https://github.com/jvzonlab/OrganoidTracker). The models used are available on Zenodo (DOI: 10.5281/zenodo.13912685, 10.5281/zenodo.13946119).

3.5 Supplementary figures

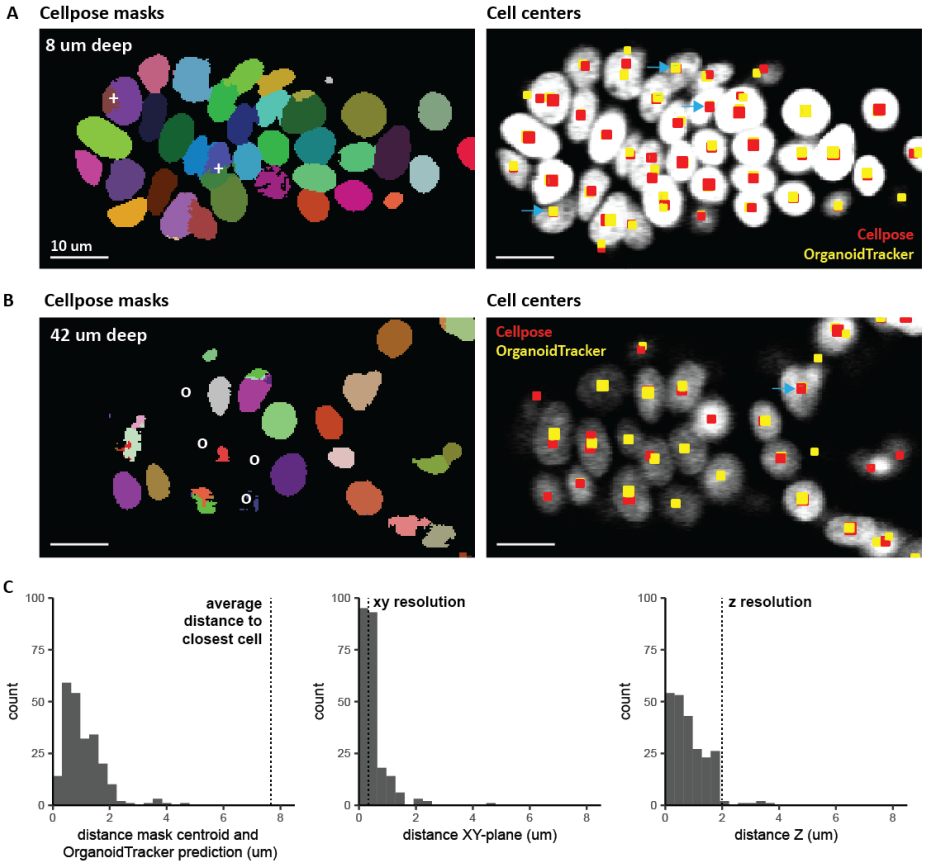


Supplementary figure 3.1: Overview of the data. Images from a representative intestinal organoid used for training our neural networks. Z-planes are chosen relative to the position of the lowest nuclei. Nuclei are labeled using a H2B-mCherry construct. Nuclei are densely packed and can have strongly elongated shapes. Signal-to-noise levels are poor deep in the organoid and late in the experiment. The z-resolution (2 µm) is much lower than the xy-resolution (0.32 µm).

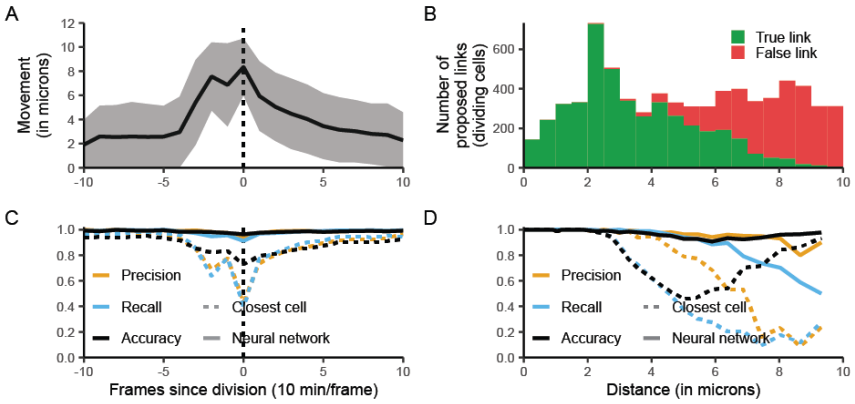


Supplementary figure 3.2: Adaptive distance mapping during cell center detection. A) Distance maps for our adaptive method (top) versus the conventional way (bottom). The standard deviation of the Gaussian mapping of distances to intensity values are indicated. They are shown in terms of pixels (px) and real distances (µm). For clarity

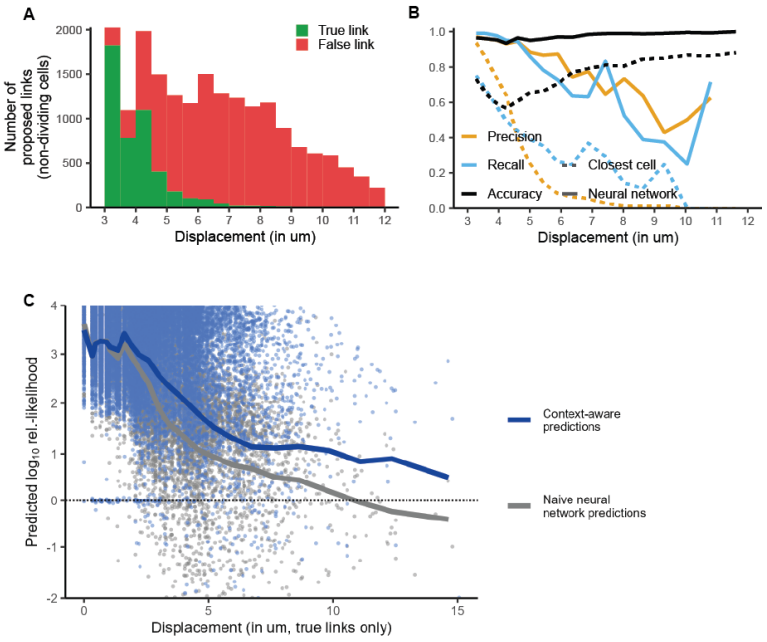
we show training data generated from the same time point and chose overlapping crops to aid comparison. Shown are the input (first column), the desired outcome (second column) and the weight assigned to every pixel during the calculation of the loss (third column). Our adaptive method allows the Gaussians around the cell centers to be much broader without overlapping, when in the conventional method Gaussians have to be narrower to not overlap. This is especially clear for non-spherical nuclei, where the centers can be very close together (see distance in red). The weights (third column) encode the contribution of a pixel towards the loss function. They are used to amplify the attention the neural network will pay to the nuclei containing regions relatively to the much larger background region. Gaussians for the weights can be chosen to be relatively broad to increase weights associated with the regions between nuclei, to ensure spots are well-separated. The edge regions, where nuclei are partially out of the field of view, receive zero weight. **B)** The performance of the adapted and non-adapted method on unseen test data. Green squares indicate predicted cell centers, smaller squares denote cell centers that lie above or below the z-slice shown. Examples are chosen to illustrate the improvements quantified in C). In the top row, the standard method over-segments a large nucleus present in many z-slices, associating it with two Gaussian spots instead of single large one. In the bottom row, the Gaussian spots produced by the standard method overlap, due to the closely packed and contorted nuclei, but is correctly separated in the adapted method. **C)** The false positive and negative rates for 5 different organoids using the adapted and the standard distance mapping during training. Organoids were weighted by the amount of cell detections. Also shown is the fold chance improvement by using the adaptive distance mapping in false positive and negative rates for 5 different organoids. Error bars denote the standard deviation. The adaptive distance mapping leads to significantly improved accuracy.



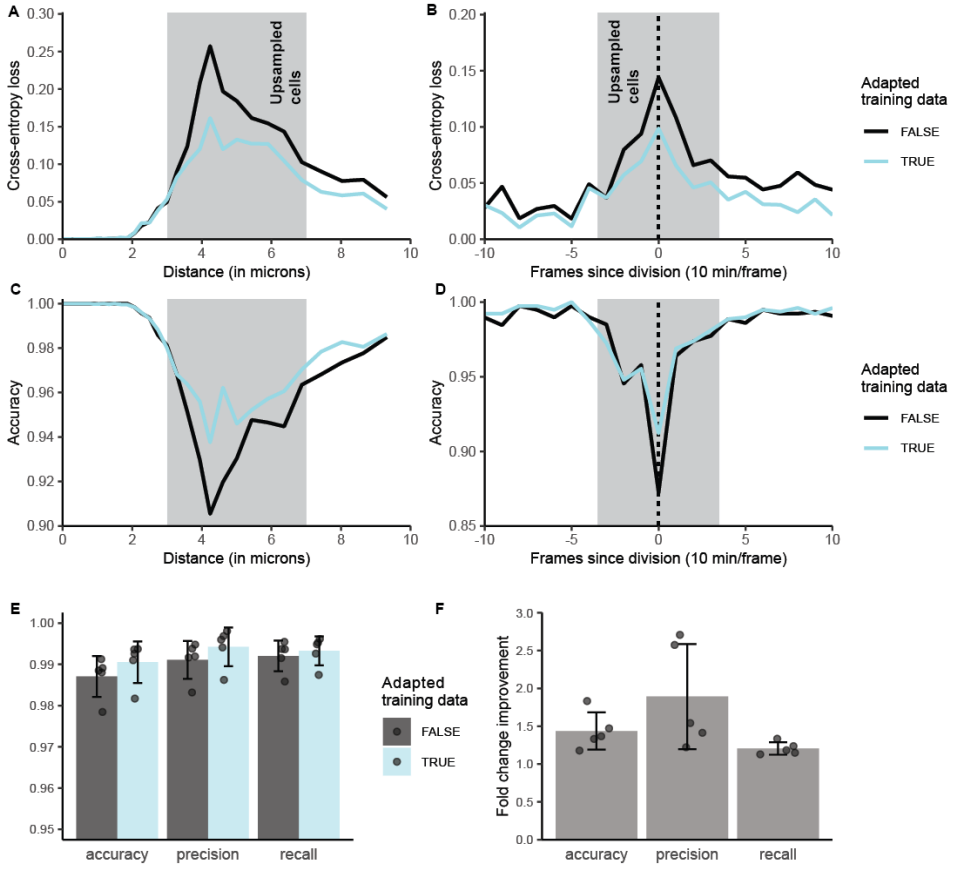
Supplementary figure 3.3: Accuracy of centroid prediction: (A) CellPose masks for the intestinal organoid data (left). Instances of CellPose oversegmentation are indicated (+). CellPose-derived centroids, given as the center of mass of the segmented 3D mask, shown in red, predicted OrganoidTracker centroids in yellow. Smaller squares are located above or below the image plane shown. The overlapping squares show the strong alignment of centroid position for the two methods, in the absence of Cellpose oversegmentation. Blue arrows indicate instances where the centroids appear off-center, but are actually centered when considering the full 3D mask. **(B)** Same as A) but at higher depth, where Cellpose performance is worse. Instances of undersegmentation are indicated (o). Centroids align well when detected by Cellpose. **(C)** Distribution of 3D distances between validated CellPose mask-derived centroids and OrganoidTracker-predicted centroids (left). Predicted centroids are much closer to each other than the average 3D neighbor distance (dotted line), ensuring that uncertainty in centroid predictions does not cause erroneous swaps in cell identity. Distances in the XY-plane (middle) are smaller and on the order of the XY-resolution (dotted line). Distances in the Z plane are higher (right), but within the lower Z-resolution (dotted), indicating that the difference in Z-position typically represents not more than one Z-plane.



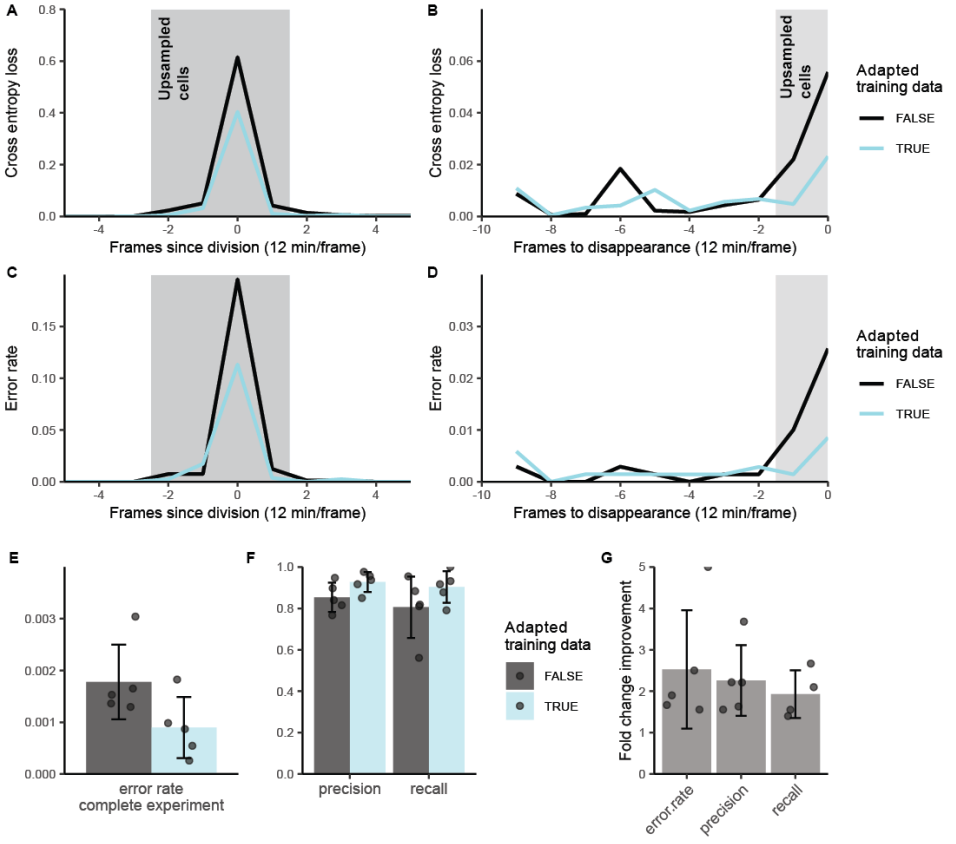
Supplementary figure 3.4: Link prediction for cells undergoing division. A) The movement of nuclei per frame plotted against the frames since (or before) division (shaded region denotes standard deviation). Around the division, nuclei move substantially faster. **B)** Distribution of travel distance per frame for cells that are within 3 frames from the moment of division. All the proposed links in the graph are shown, whether true (green) or false (red). The distribution for true and false proposed links shows large overlap for dividing cells. C) Precision, recall and accuracy are shown as a function of frames since (or before) division. Naively picking the closest cell (dotted lines) breaks down as a mechanism for predicting links around division, when less than half of the true links is recalled (blue dotted line). In contrast, the neural network (solid lines) shows highly accurate prediction. D) Precision, recall and accuracy for the dividing subset of cells as a function of distance travelled by the nuclei. The neural network (solid lines) stays accurate (black line) even for fast movements, while linking the closest nuclei breaks down for dividing cell moving >3 μm in a single frame.



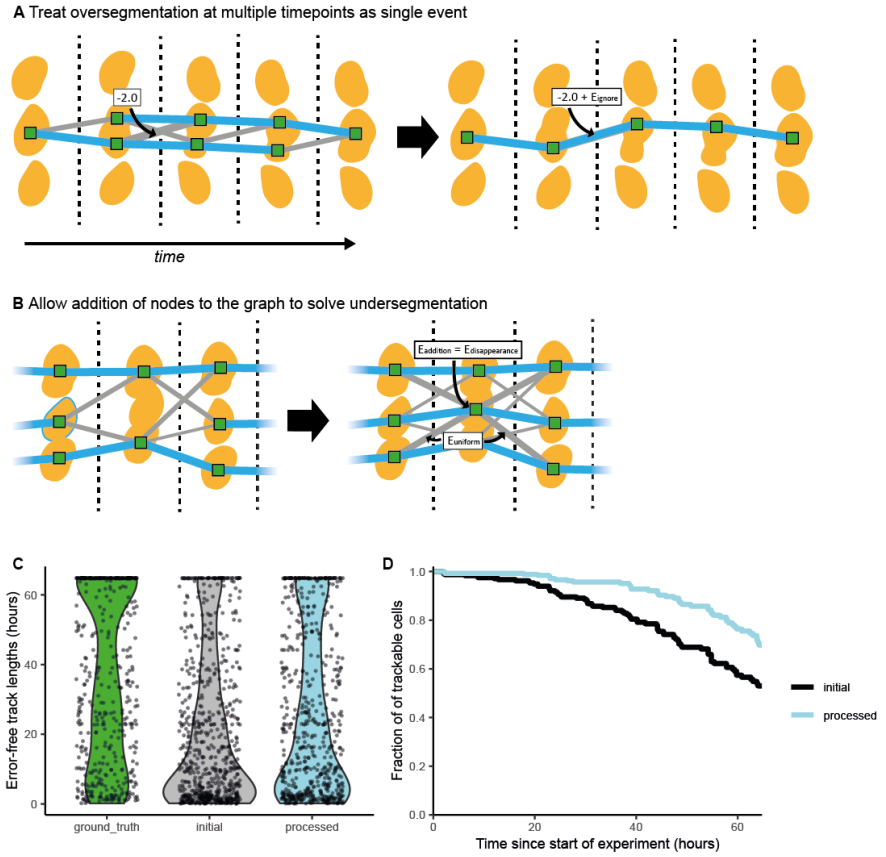
Supplementary figure 3.5: Performance for non-dividing fast moving cells. A) Characterization of non-dividing fast-moving cells (>3 µm per frame) based on manually curated data shows a tail of large displacements up to 7 µm. **(B)** The precision, recall and accuracy of the link prediction neural network (solid lines) for non-dividing fast-moving cells (>3 µm per frame), compared to predictions based on the 'smallest displacement' criterion (closest cell, dotted lines). Neural network accuracy is high over the full range, with recall and precision remaining high even for displacements larger than the typical cell diameter (~5 µm). **C)** The effect of integrating context in the link predictions (plotted as the relative log-likelihood of being true) for true links plotted against the displacement. Dots are individual link predictions before (gray) and after introducing context (blue). Lines represent moving averages. The difference between the naive and context-aware predictions increases with the displacement to a final improvement of more than an order of magnitude, showing that marginalization is especially beneficial in resolving the uncertainty stemming from large displacements.



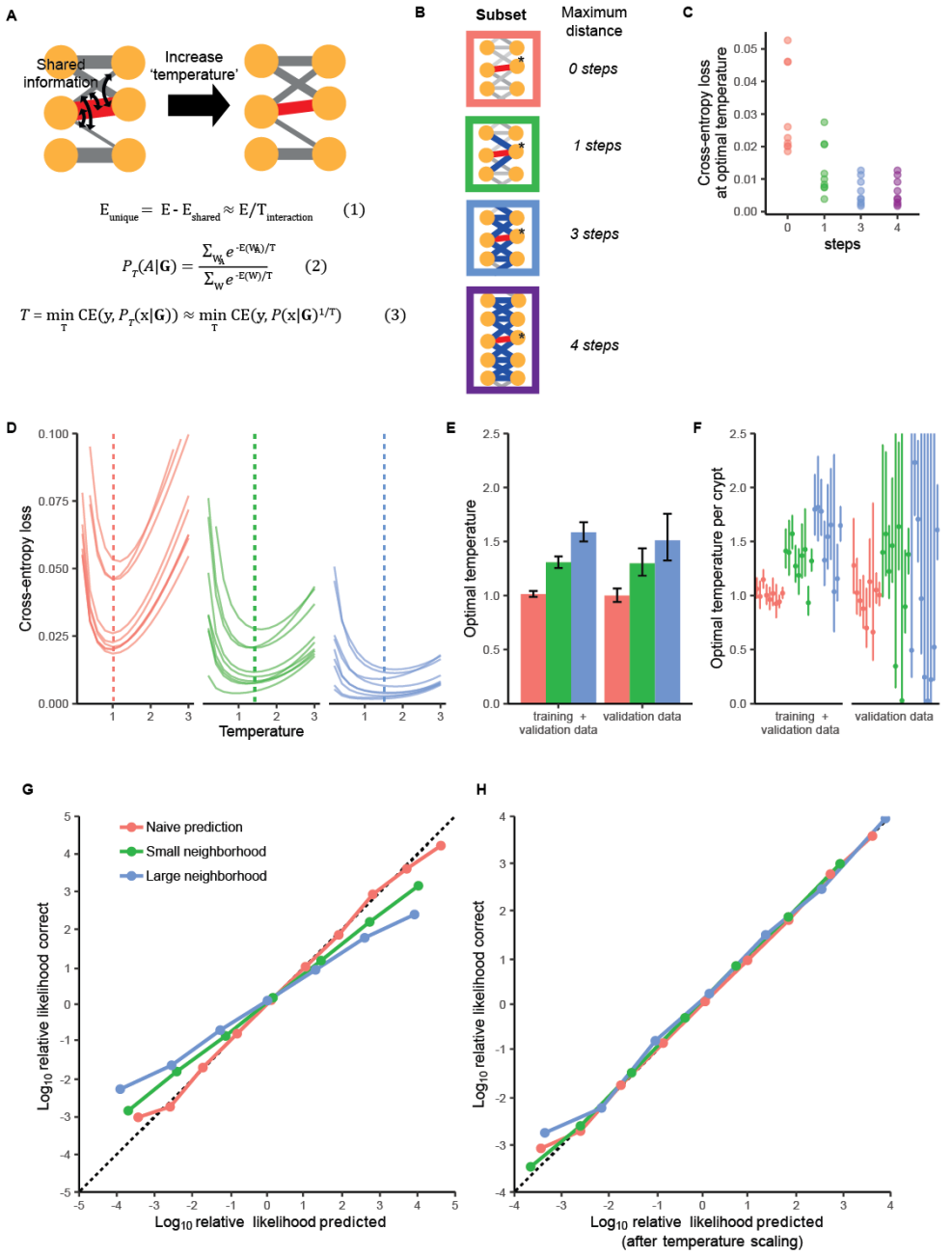
Supplementary figure 3.6: Impact of up-sampling the training dataset for link prediction. A, C) The cross-entropy loss and accuracy for link prediction as a function of the distance between the two cell detections in the proposed link, for an up-sampled (blue) and regular training dataset (black). In the up-sampled training data challenging links with a distance of between 3 and 7 um (gray region) were increased by a factor of 5. Even small improvements in the cross-entropy loss and accuracy can lead to substantial improvements in tracking. The flow-solver that finds optimal tracks optimizes over all links, so improving the classification of a single link will help tracking on all nearby cells. Similarly, the error probabilities are calculated by integrating information from many links and thereby combine all the accuracy gains. **B, D)** The cross-entropy loss and accuracy for link prediction as a function of the time since (or before) division of the cell that forms the source node of the link. A five-fold up-sampling in the training data (blue) of cells close to division increases accuracy around the moment of division. **E, F)** Precision and recall before and after up-sampling the training data, for the complete dataset. Error bars shows the standard deviation of the 5 individual organoids that make up the test data. **G)** Fold change improvement in accuracy, precision and recall for every organoid in the test data upon up-sampling the training data. Error bars denote the standard deviation between different organoids.



Supplementary figure 3.7: Impact of up-sampling the training dataset for division prediction. A, C) The cross-entropy loss and error rate for cell division prediction around the moment of division for an up-sampled (blue) and regular training dataset (black). The up-sampled training data increased cells around division (gray region) 10-fold. **B, D)** The cross-entropy loss and error rate for cell division prediction as function of time prior to the end of a cell track. The majority of cell tracks end due to cell death. As dying cells resemble dividing cells, the error rate without up-sampling increases to almost 3% (black line). For up-sampled training data (blue line), the error rate only increases for the final time point before death and only to 1%. **E)** Error rates for all cell detections in the test data with (blue) and without up-sampling the training data (gray). Error bars shows the standard deviation of the 5 individual organoids that make up the test data. **F)** Precision and recall measures with and without up-sampling the training data. **G)** Fold change improvement in error rate, precision and recall for every organoid in the test data upon up-sampling the training data. Error bars denote the standard deviation between different organoids. Both the amount of false positives (precision) and false negatives (recall) decreases around two-fold by up-sampling the training data.

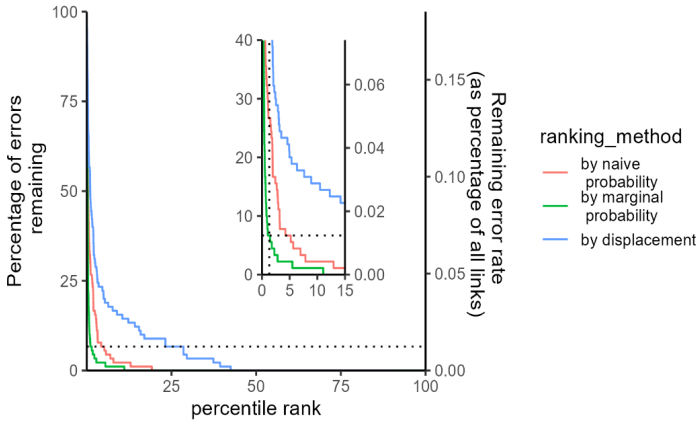


Supplementary figure 3.8: Post-processing of the maximum likelihood tracking solution. A) Illustration of how post-processing deals with over-segmentations. The blue links represent the maximum-likelihood solution, while the grey links are excluded from the solution. Over-segmentations are not independent events, but are treated as such by the min-cost flow solver, which can lead to fragmentation of tracks (left panel). During post-processing we check if track fragments can be joined by adding a single link and removing the cell detections associated with the over-segmentation. Removing multiple of these nodes is associated with a single penalty (E_{ignore}, the relative log-likelihood of over-segmentation) which is added to the new link so that our probabilistic description remains correct. **B)** Illustration of how post-processing deals with under-segmentations. The min-cost flow solver cannot add nodes to the graph, which can lead to fragmentation of tracks (left panel). During post-processing we check if track fragments can be joined by adding a node. Adding a node is associated with a penalty ($E_{addition}$, the relative log-likelihood of missing a cell detection) and the node is connected to nearby nodes in the graph with uniform probability. **C)** Error-free track length distributions in a representative tracked organoid. The ground truth (green) represents fully manually corrected data, where tracks are only cut short by cell death, the end of the experiment, or leaving the field-of-view. The initial maximum-likelihood solution (gray) has many more short tracks, which is partly solved by post-processing (blue). **D)** The fraction of cells present at a certain time that can be tracked without error from the start of the experiment. Only cells that are trackable for the full experiment in the ground truth are considered. By post-processing the data we can increase the fraction of cells that are trackable for the complete experiment (> 60 hours) from around 50% to around 75%. See the section 'Solving over- and undersegmentation' in the Methods for more details.

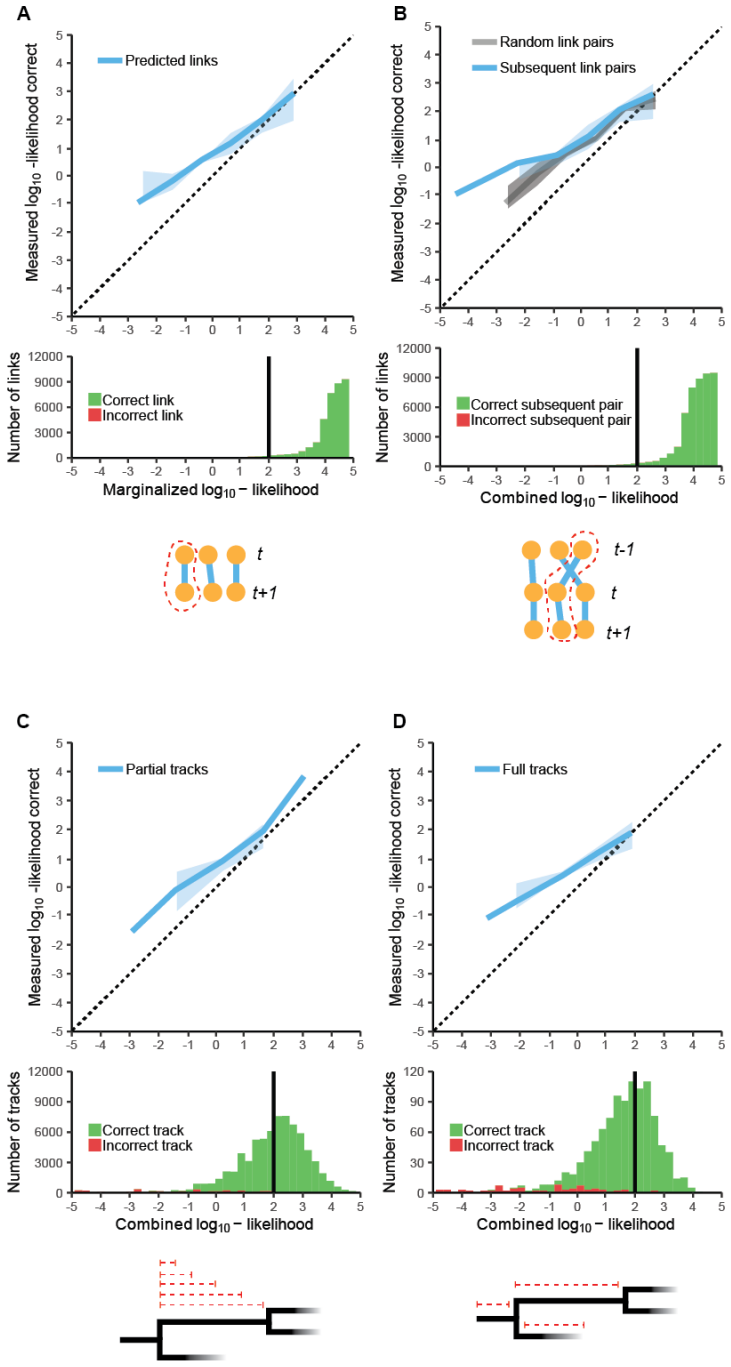


Supplementary figure 3.9: Temperature scaling during the marginalization procedure. A) Visual illustration of the temperature scaling procedure. The prediction on a single link (encoded in its weight) is not independent from predictions on nearby links. To account for this shared information we divide the link energy by a 'calibration temperature' (Eq. 1). The marginalized link probability, P(A|B), is then based on scaled energies (Eq. 2). The calibration temperature is found by minimizing the cross-entropy loss (CE) between the marginalized predictions and the ground truth (y), with respect to this temperature (Eq. 3). This scaling temperature has to be calibrated only

once after training a set of division and link neural networks. The calibration can be done on the same data that was used for neural network training. For data far outside the training distribution new calibration can be performed on a small manually corrected set of links. **B)** Illustration of the difference subsets used in the subsequent plots. Red symbolizes the naïve approach, where the subset simply is the link of interest. In this case no marginalization is done. The green subset only considers link at a distance of one step from the target node of the link of interest. The blue subset goes up to a distance of three steps. This is the largest set that is computationally feasible (~ 1 hour of computation time for 300 frames). **C)** Prediction performance as measured by the binary cross entropy loss versus the subgraph size, all dots are individual organoids. The loss at the optimal temperature (lowest point in panel D) was taken. Increasing subgraph size improves prediction, but there is no improvement when going beyond three steps, see method section 'Marginalization' for further discussion. **D)** Cross entropy loss between predictions and ground truth (based on all 9 organoids in the training dataset) for different neighborhoods and temperatures. Minimum loss (dotted line) is achieved at higher temperatures when the subset gets larger. All lines represent individual organoids. **E)** Optimal calibration temperature for every neighborhood. The error bars represent the 95% confidence interval (this is a lower bound as not all observed links are truly independent). Limiting ourselves to calibration on the validation dataset that was left during training gives the same results. **F)** Optimal temperatures and confidence intervals (again lower bounds) per organoid. The estimates show great overlap between organoids, validating that we can use a single calibration temperature for all. **G)** Predicted versus actual log-likelihoods after marginalization on different subsets. When marginalizing on larger subsets predictions become overconfident. If predictions suggest that links are not correct, more than expected fraction is actually correct and the other way around. **H)** After scaling the energies with the proper temperature per subset, we get well-calibrated link predictions.

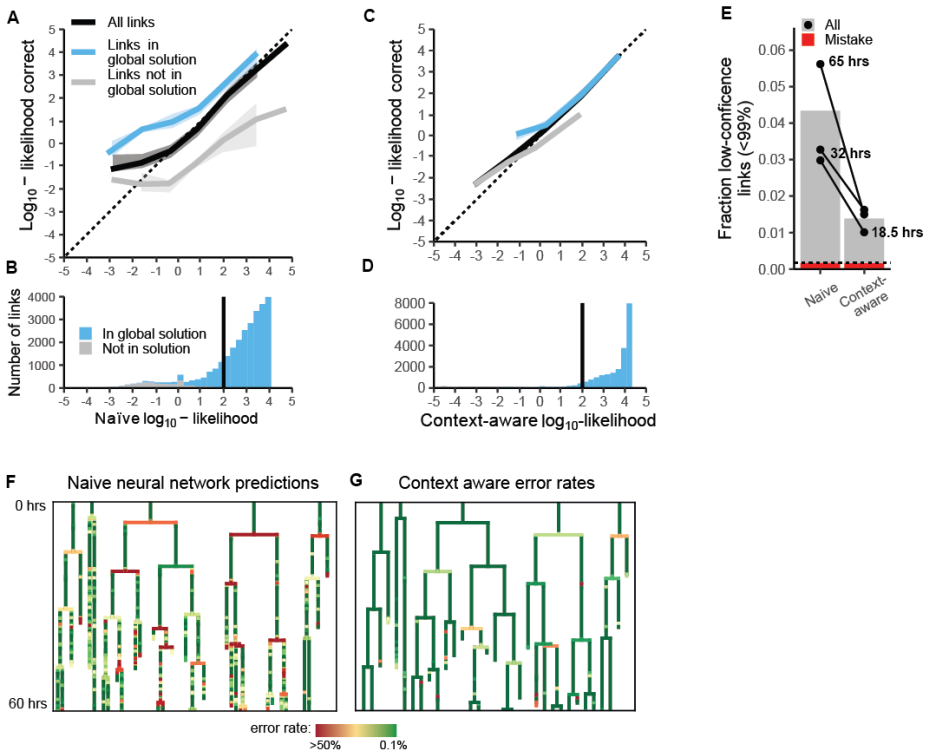


Supplementary figure 3.10: Error rates versus heuristic. The comparative performance of marginalized error rates against naïve error rates and a typical heuristic for linking errors, distance travelled, is shown. If all links with an error rate of above 1% are reviewed, >90% of the errors in the data are corrected (horizontal dotted line). This leads to a remaining linking error rate of at most ~ 0.01% per cell per time point. This is an upper bound as review might also correct nearby errors that above the confidence threshold. To achieve this, the user has to check only the ~1% unlikeliest links in the data (vertical dotted line in inset). Using the naive neural network predictions, this would require checking ~5% of the data (red curve in inset). Without probabilistic information, a user would have to manually check the ~25% biggest displacements in the dataset to reach comparable accuracy (blue curve).

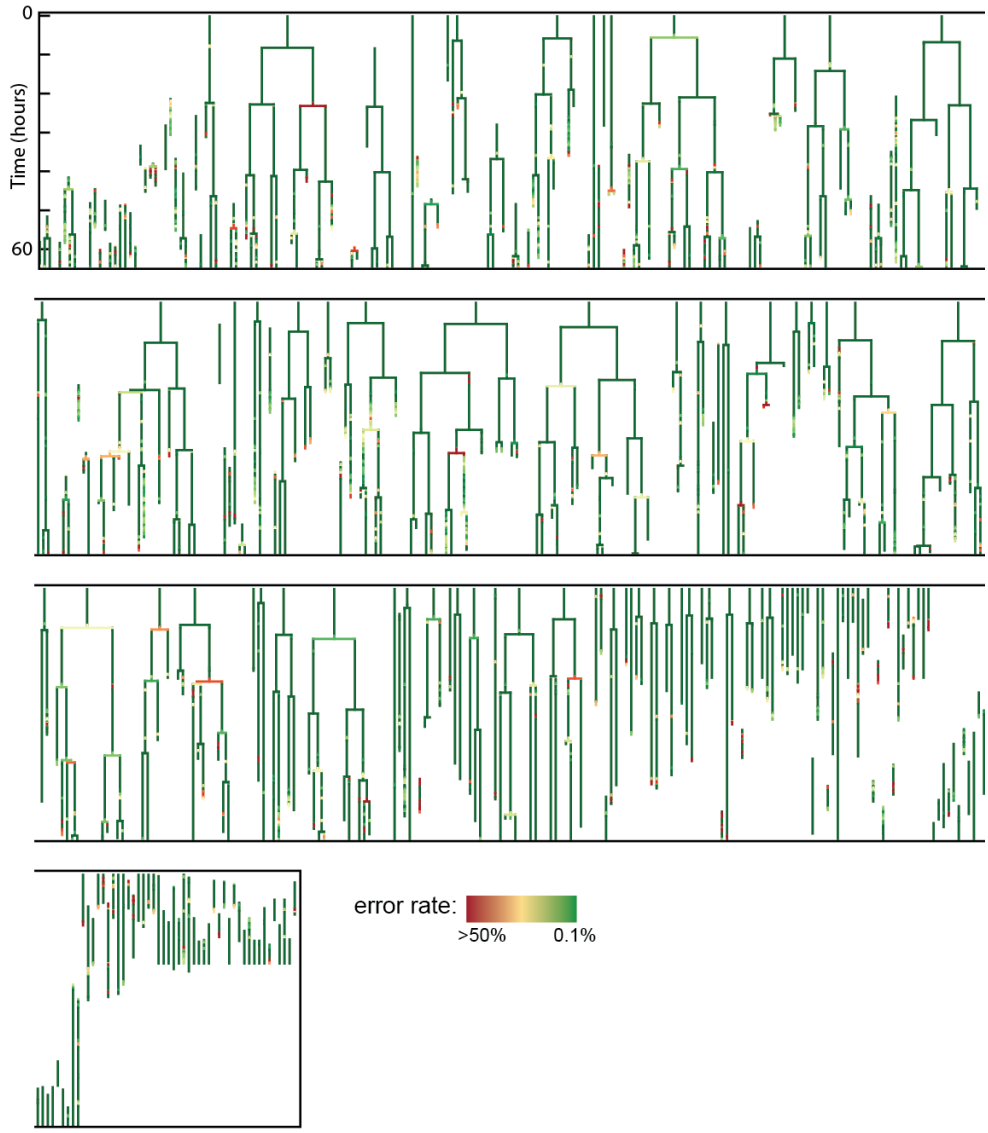


Supplementary figure 3.11: Track error rates. A) The predicted likelihood (after marginalization) of links in the tracking solution against the measured likelihood of being correct based on the manual data (top panel). The line does not overlap the dotted line, because not all graph information available to the flow-solver can be integrated in

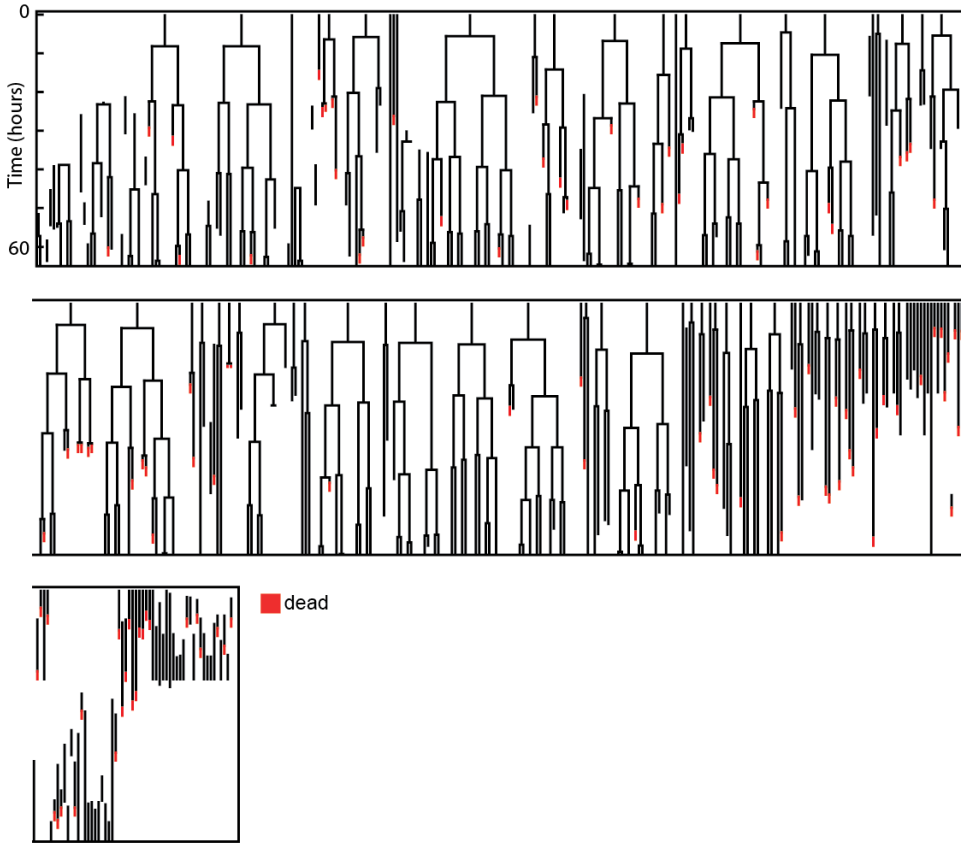
the error prediction. Key is that they remain above the dotted line so that the error predictions are conservative. The histogram (bottom panel) shows the distribution of likelihoods of correct (green) and incorrect links (red). Because the linking error rate is so low, incorrect instances cannot be seen in the histograms A) and B). The black vertical line indicates the threshold of 99% chance of being correct. **B)** The likelihood of a pair of links both being correct can be calculated by combining their constituent probabilities by simple multiplication. It does not matter if the links are subsequent (blue line) or unconnected (gray line). It is thus not so that a link being true is informative of the truth of the subsequent link, beyond its predicted error probability. **C)** The predicted error rates for tracks of arbitrary length. The probability that a cell can be correctly traced back to its last division (red dotted lines) is predicted for every cell at every timepoint. The probability of the track being correct is calculated by multiplying all the constituent probabilities. The tracks are compared to the ground truth and deemed correct if they recapitulate it exactly, yielding a similar calibration curve to A). **D)** The same as in C) but now only with tracks that span the full cell cycle, again producing a similar calibration curve as in A). Error detection works efficiently for full tracks spanning the complete cell cycle, only one incorrect track is above the 99% certainty cut-off.



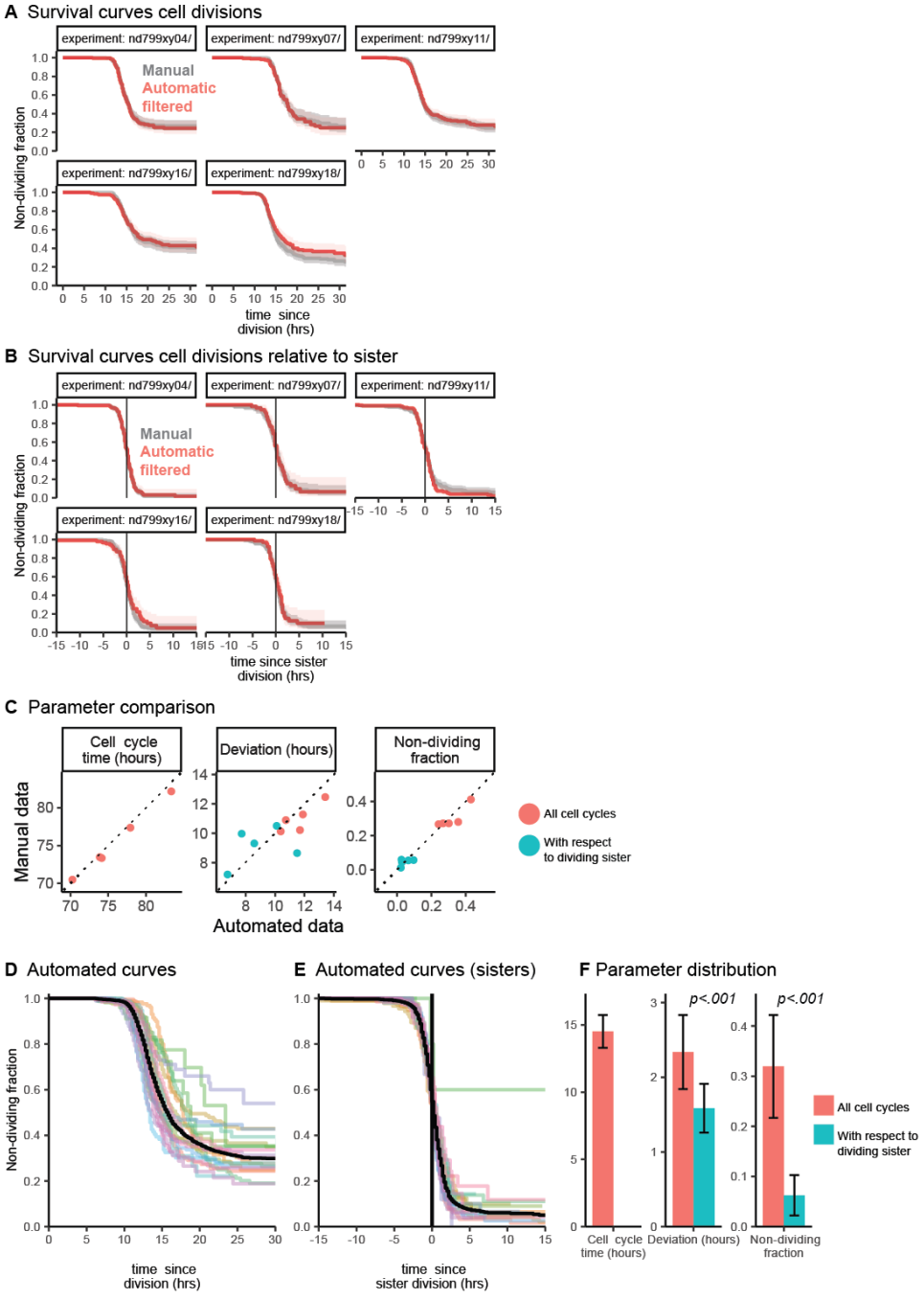
Supplementary figure 3.12: Marginalization during full procedure. A) The predicted relative log10-likehoods from the link neural network (black line) is well-calibrated when compared to the manually corrected ground truth (dotted line denotes perfect calibration). The overestimation at low likelihoods is due to the filtering of unlikely links before tracking and marginalization (see Methods). This filtering preferentially removes low-probability incorrect links, making the remaining ones more likely to be true. The fact that links in the global solution (blue line) are much more likely than expected and links excluded from the tracking solution (gray line) are less likely than expected, suggests that contextual information could improve the error rate prediction. **B)** Many links in the tracking solution are less than 99% (black line) certain based on the naïve predictions. **C)** Marginalization integrates context and largely removes the discrepancy between links in and out of the tracking solution. **D)** Only few links in the tracking solution are less than 99% certain after marginalization. **E)** The fraction of uncertain links (<99% certainty) as fraction of total. Marginalization decreases the number of uncertain links around four-fold. The longest experiments, which have poorer signal to noise, benefitted most from the marginalization (the black lines denote individual experiments). The red fraction indicates actual errors in the low-confident fraction. The dotted line the fraction of errors across all links including ones that have a high probability of being true, showing that no significant number of errors is missed. **F)** Five randomly selected lineage trees colored by their naïve predicted error rate (yellow to red links are uncertain. **G)** The same trees after marginalization.



Supplementary figure 3.13: All trees with error rates before correction for typical organoid. All cell tracks over 5 hours long for a single organoid with one crypt plus villus region imaged. Data is shown before any correction and with the context-aware error rates. Note that many cells can already be traced back to the start going only through high-confidence (green) links.

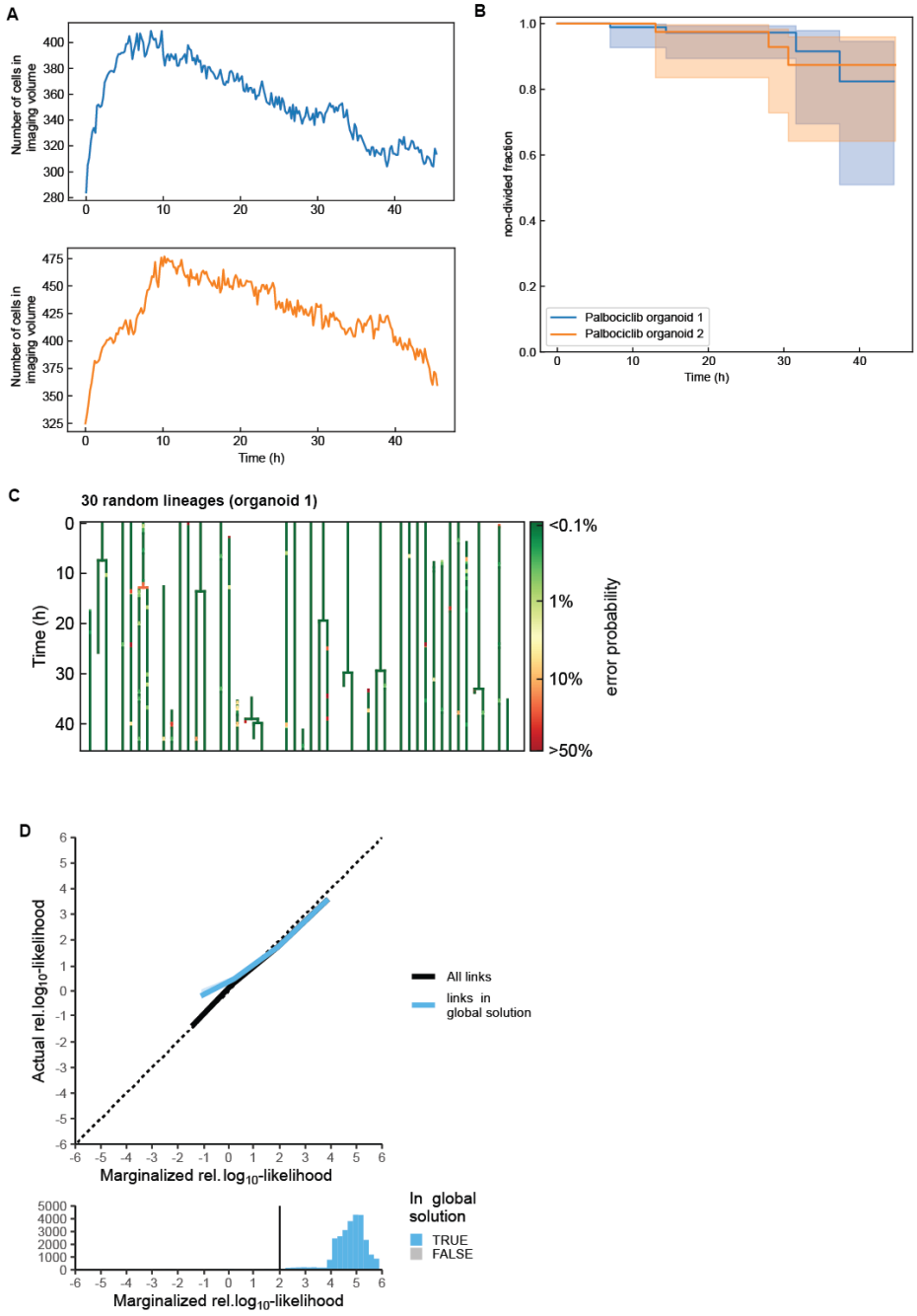


Supplementary figure 3.14: All trees with error rates after correction for typical organoid. All cell tracks over 5 hours long for a single organoid with one crypt plus villus region imaged. All track that end and are not marked 'dead' leave the field of view.



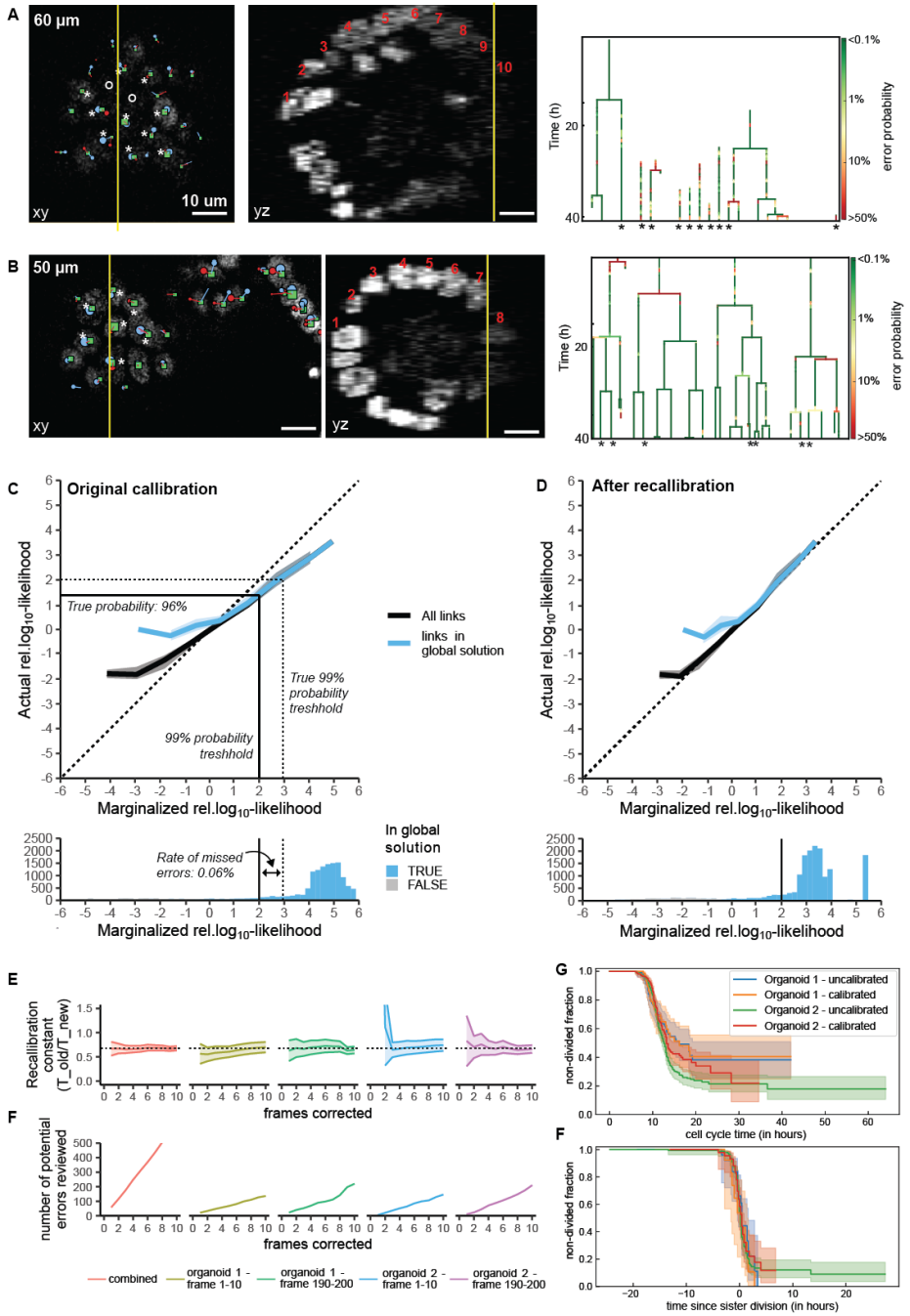
Supplementary figure 3.15: Survival curves for different organoids. A) Kaplan-Meier curves describing the fraction of non-divided cells as a function of time since division, for 5 organoids where manual reference data was available for the complete crypt (grey curve). The shaded region denotes the 95% confidence interval. **B)** Kaplan-Meier curves describing the fraction of non-divided cells as a function of the time since the division of the sister

cell. **C)** Comparison of manually annotated data to automated analysis for the estimation of three key parameters of lineage dynamics. Color indicates whether the variation in cell cycle duration or the probability that the cell will divide again was calculated for all cell cycles (red) or for cell cycles relative to the moment of division of the sister cell (blue). Cell cycle times show almost perfect correlation between manual and automated data. The deviation in the cell cycle time is more sensitive to outliers and consequently shows poorer correlation. The non-dividing fraction again shows strong correlation. **D)** Overlay of all Kaplan-Meier curves of automatically tracked organoids, n=20, which describes the fraction of non-divided cells as a function of time since division. Black line represents the combined data. **E)** Same but as a function of time since the sister division. **F)** Statistical analysis of organoid lineage parameters. Error bars indicate standard deviations around the mean. Cell cycle times only show limited variation across organoids and experiments, with standard deviation <10% of the mean cell cycle time (left panel). The variation around the cell cycle mean for cell cycle times is significantly larger than the variation between the cell cycle times of sisters (middle panel), suggesting strong correlation between sister pairs. This is further supported by the results that given that the sister divides, the non-dividing fraction becomes close to zero instead of around 30% (right panel).



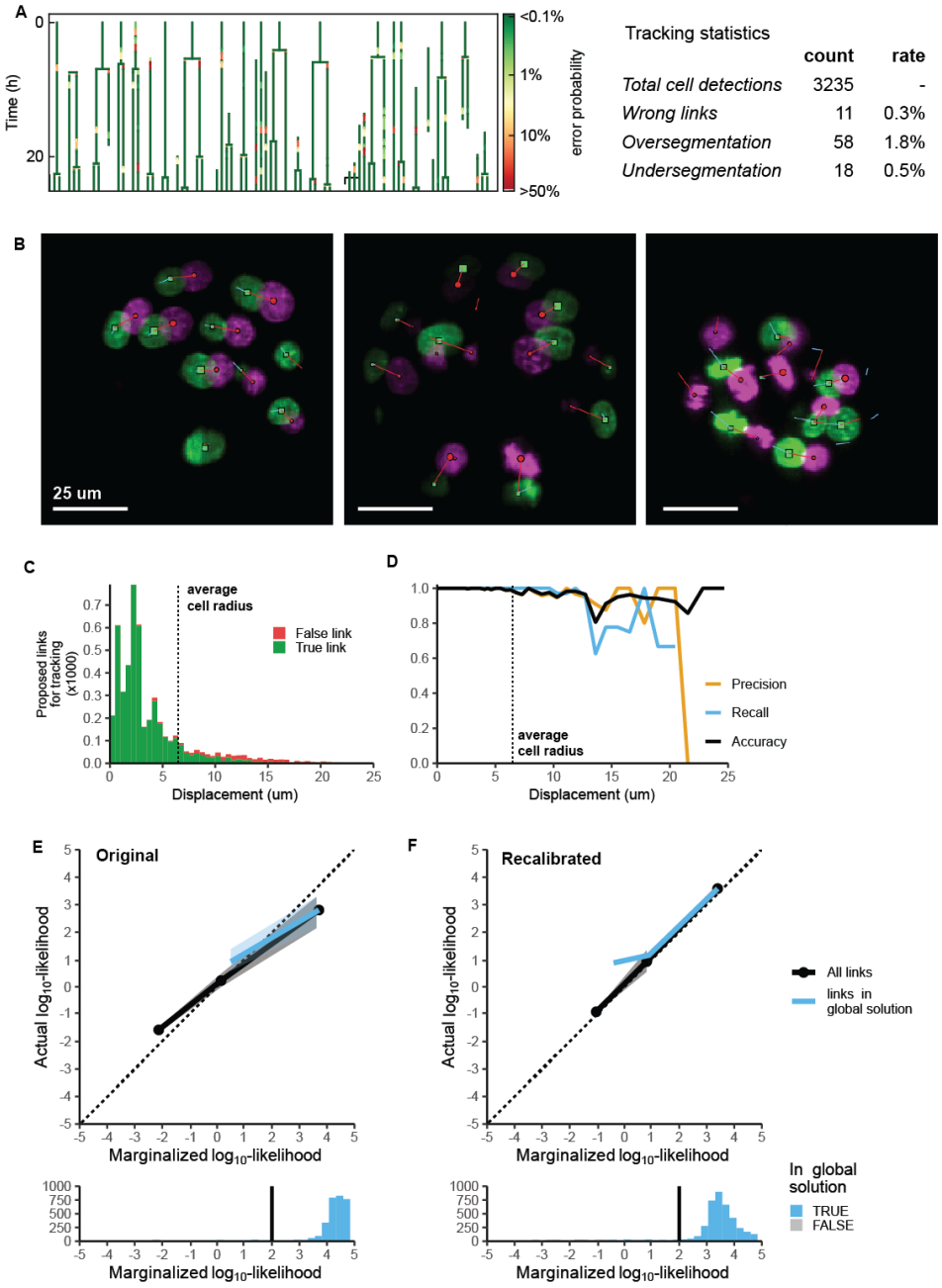
Supplementary figure 3.16: Tracking performance on organoids with perturbed dynamics. (A) The number of detected cells in the imaging volume for two organoids treated with palbociclib, which blocks cell division. After ~10 hours of Palbociclib exposure, cell death dominates over cell division, leading to decreased cell numbers. **(B)** Distribution of cell cycle lengths (0 hours is cell cycle start) plotted for the same two organoids. Cells divide either slowly (>25 hours) or not at all, consistent with Palbociclib action. **(C)** Thirty randomly selected lineages show

tracking with very low potential error rates. Cells generally divide only once, or not at all. **(D)** Top: measured link likelihood versus likelihoods predicted by the neural network. Data is shown for all possible links (black) or links that are either in the global solution (blue). Dotted line is perfect calibration. Data is for the 2 different organoids in (A) and (B). Shaded region is S.D.M. Predictions for Palbociclib-treated organoids remain well-calibrated, indicating that perturbation of cell dynamics does not affect the quality of error predictions. Bottom: histogram of the predictions. Almost all links in the tracking solution (blue) are above the 99% probability threshold (vertical line) indicating highly confident tracking.



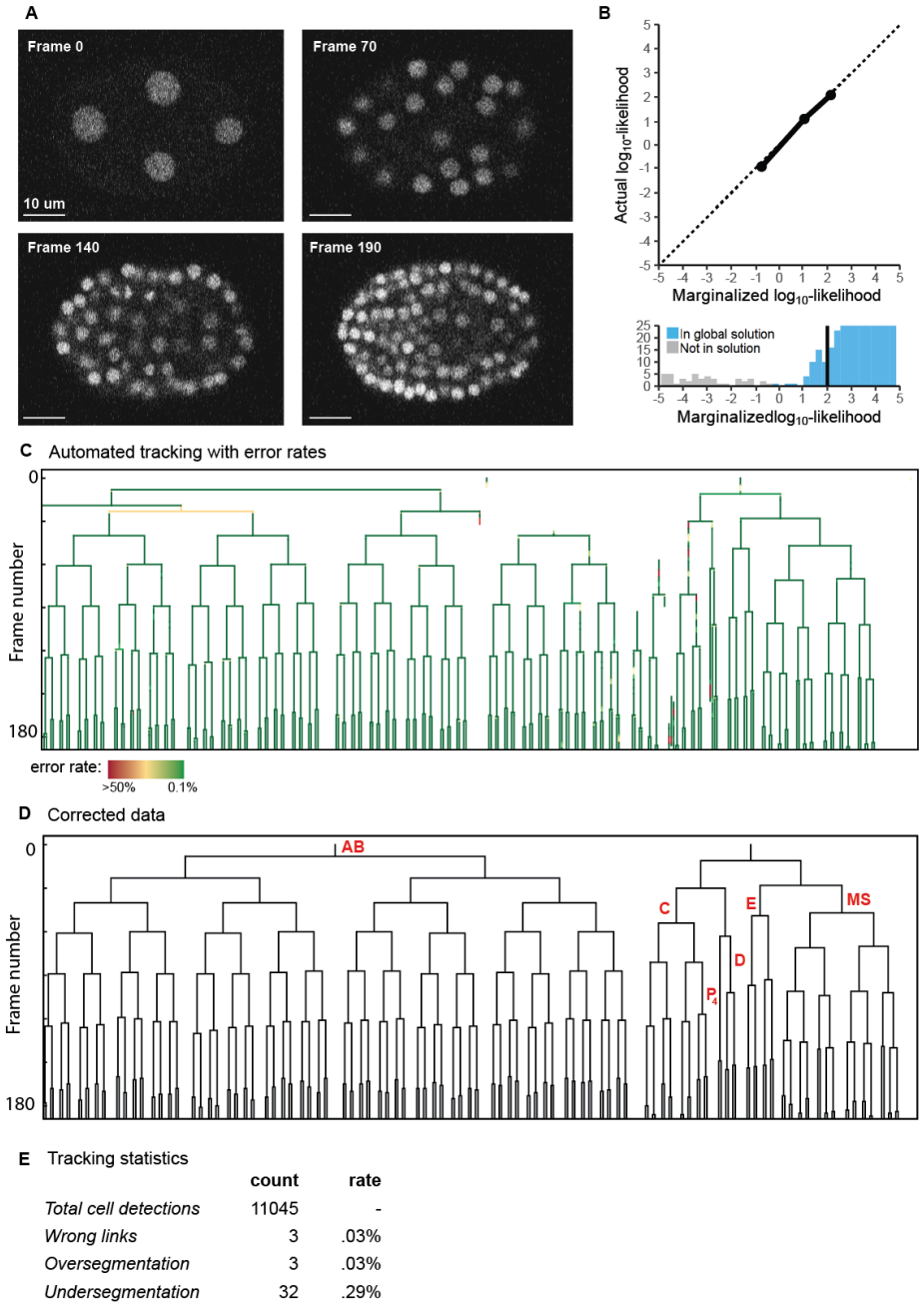
Supplementary figure 3.17: Performance for different image acquisition parameters. (A) Tracking performance for intestinal organoids imaged on a different scanning confocal microscope (Leica TCS SP8) than used for collecting the training data (Nikon A1R MP). Consequently, imaging data had lower planar resolution (0.4 µm/px rather than 0.32 µm/px), cells could be imaged deeper into the organoid, but at low signal-to-noise. Green

squares are tracked cells at 60 µm depth, while red and blue squares are their previous and next locations, respectively. The cells marked with an asterisk (*) correspond to cells in the lineage trees (right). Open circles (o) denote missed cell detections. Even at this comparatively low signal-to-noise ratio most (>80%) cells are detected. The YZ-cross-section show the depth of these cells in terms of cell number. Lineage trees are color-coded by predicted error rate. Most cells can be tracked for multiple hours without potential mistakes. **B)** Same as in (A), but for cells at 50 µm. Signal-to-noise ratios are higher here and all cells are detected, while lineages show cell tracking for long (>10 hrs) periods without potential mistakes. **(C), (D)** Top: measured link likelihood versus likelihoods predicted by the marginalization procedure either without (C) or with (D) recalibration on newly corrected data. Data is shown for all possible links (black) or links that are either in the global solution (blue). Dotted line is perfect calibration. Data is for the 4 different data sets shown in (E) and (F). Shaded region is S.D.M. Without recalibration (C), predicted likelihoods are overconfident, e.g. the 99% probability threshold actually corresponds to a lower level of certainty. Using the actual 99% probability threshold would increase the number of links needed to be checked (C, bottom histogram). However, the number of flagged potential mistakes omitted when using the uncalibrated threshold remains low (0.06% of links). **E)** Visualization of the recalibration process. The recalibration constant, defined as the ratio of the old and new scaling temperature, is estimated by manual error correction of predicted tracks. Graphs show the estimated recalibration constant versus the number of frames corrected for four different data sets, as well as all data pooled. The estimates converge on the consensus estimate (dotted line) for >5 corrected frames. Shaded area denotes 95% confidence interval. **F)** The number of reviewed potential errors (<99% certainty links) as a function of frames corrected. Tight estimates of the recalibration constant are reached after ~200 reviewed potential errors. **G)** Survival curve indicating the probability that a cell has not divided at time t after its birth. Shown are the original (blue, green) and recalibrated (orange, red) data for the same organoid. **H)** Same as F) but for the timing of division relative to the sister division. The strong overlap between original and recalibrated data indicates that for downstream applications perfect calibration is often not essential.



Supplementary figure 3.18: Performance on different system: blastocyst. (A) Lineages color-coded by error-rates for a blastocyst developing from the 16 to 64-cell stage, as obtained by tracking without manual curation. Many cells can be tracked from the beginning to the end without any potential mistakes. **B)** Blastocyst cells undergoing large movements, with subsequent timepoints shown in green and magenta. Green square and red circles indicate the current and subsequent detected centroid position, respectively. The red line connects these positions, while the

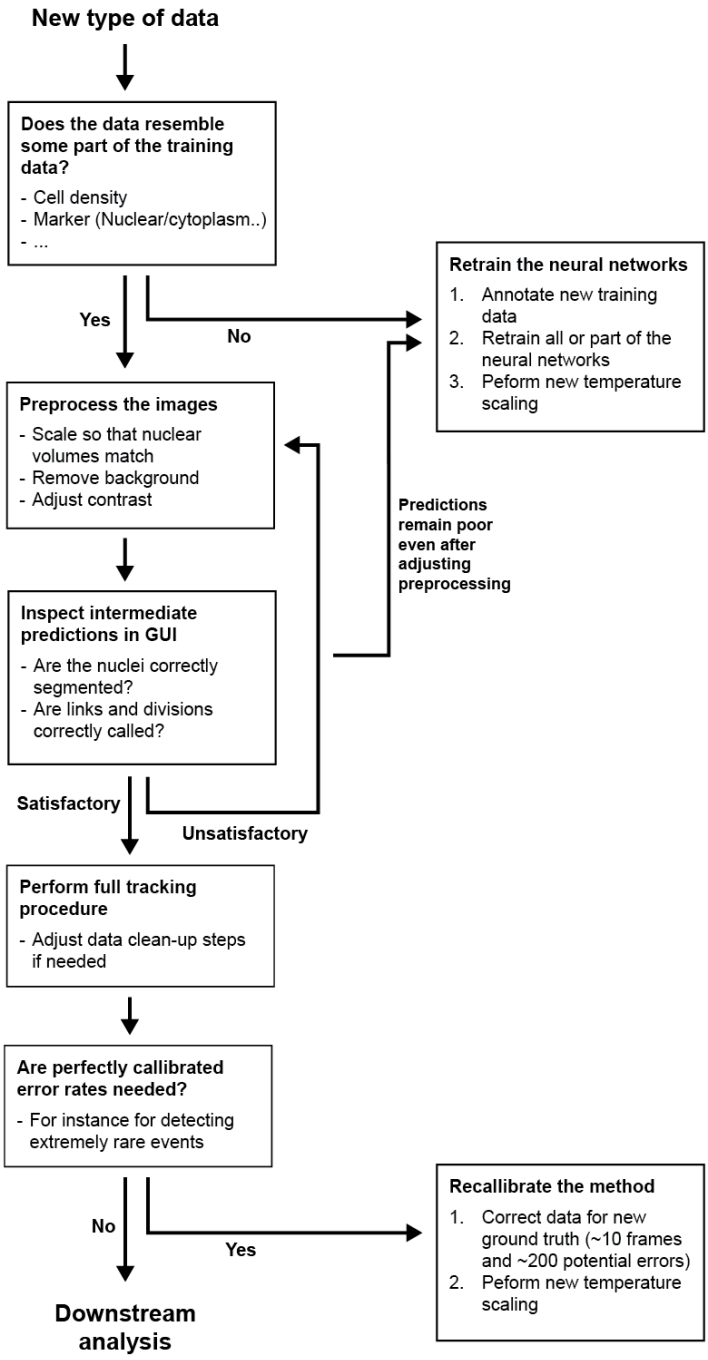
blue line connects to the centroid position in the previous time point. When either marker is not present, then the corresponding positions are >4µm away from the z-slice shown. Our method tracks cells even for displacements larger than the typical nucleus size. **C)** Histogram of potential cell displacements in a single frame, for all potential links, with the correctness of each link established by manual curation. Many true links represent fast-moving cells, corresponding to displacements between timesteps of >1 cell radius, or ~6µm, large compared to cell movements in intestinal organoids both in absolute and relative terms. **D)** Precision, recall and accuracy versus displacement between timesteps, for the optimal tracking solution before manual curation. Even fast-moving cells are tracked with high accuracy. **E), F)** Top: measured link likelihood versus likelihoods predicted by marginalization, either with (E) or without (F) recalibration on newly corrected data. Data is shown for all possible links (black) or links that are either in the global solution (blue). Dotted line is perfect calibration. Data is for 2 different blastocysts, corresponding to the data in (C) and (D). Shaded region is S.D.M. Without recalibration (C) predicted likelihoods are overconfident. Using a recalibrated threshold does not majorly increase the number of links to be checked (D, bottom histogram), as the bulk of links in the tracking solution remains above the 99% probability threshold (black vertical line).



	count	rate
Total cell detections	11045	-
Wrong links	3	.03%
Oversegmentation	3	.03%
Undersegmentation	32	.29%

Supplementary figure 3.19: C. elegans tracking. A) Z-slices of 3D confocal data for C. elegans embryogenesis (training data). **B)** Using retrained neural networks, we obtain well-calibrated link-likelihood predictions (top panel). Almost all links in the tracking solution that minimize the global energy (blue) have very high marginalized likelihoods (bottom panel). In this data set, only few links in the graph are not in the tracking solution (grey) as the

nuclei are less closely packed then in intestinal organoid data. **C)** Lineage trees with associated error rates show that only few potential errors (41 links with an error probability above 1%) are present after automated tracking. **D)** Lineage trees after manual correction of potential errors. The lineages map exactly on the known C. elegans AB, E, MS, C, D and P4 sub-lineages, strongly suggesting that no errors remain. **E)** Tracking statistics show a very low error rate and that almost no error in linking. Most errors arise from undetected nuclei (under-segmentation), deep in the imaging volume.



Supplementary figure 3.20: Pipeline for using already trained neural networks on new (out-of-sample) data.

3.6 Appendix: Probabilistic description

3.6.1 Full Bayesian description of link error prediction framework

We can summarize the aim of our method as predicting the probability of a given link given the predictions made by the neural network on its likelihood and the likelihood of the links that are part of the surrounding subgraph. This predicted probability is thus also dependent on the ground truth data on which the neural networks (Platt scaling) and marginalization procedure (temperature scaling) were calibrated. In the description below we provide a full Bayesian description of these dependencies starting from the 'raw' neural network predictions, that is before calibration via Platt scaling. The energies related to the neural network predictions are thus given by:

$$E = A + B * \log(\frac{p_{raw}}{1 - p_{raw}})$$

With A and B coeffients estimated during Platt scaling on the ground truth data.

Now we can write our prediction on link l_i given the predictions made on the subgraph (G_i) and the ground truth data used for calibration (D_{GT}):

$$P(l_i|G_i, D_{GT}) = \int_{T,A,B} P(l_i, T, A, B|G_i, D_{GT})$$

With T as the scaling temperature of the marginalization procedure. Here we have marginalized over the distribution of the Platt and temperature scaling parameters to produce our final probability estimate. We can now rewrite this more conveniently as:

$$P(l_i|G_i, D_{GT}) = \int_{T,A,B} P(l_i|G_i, T, A, B) * P(T, A, B|D_{GT})$$

And further as:

$$P(l_i|G_i, D_{GT}) = \int_{T,A,B} P(l_i|G_i, T, A, B) * \frac{P(D_{GT}|T, A, B)}{P(D_{GT})} * P(T, A, B)$$

In order to evaluate this integral we would have to estimate a prior for our scaling factors ($P(T, A, B)$) and estimate $P(D_{GT})$ via numerical methods. At the same time, probabilities of the type $P(D_{GT}|T, A, B)$ are likelihoods that we can easily optimize for via logistic regression (see sections on Platt scaling and temperature scaling in the methods). We thus opt for an approximation using the maximum likelihood estimates as point estimates going forward:

$$P(l_i|G_i, D_{GT}) \approx P(l_i|G_i, T = T_{max}, A = A_{max}, B = B_{max}) = \frac{\sum_{W_i} e^{-E_{max}(W_i)/T_{max}}}{\sum_{W} e^{-E_{max}(W)/T_{max}}}$$

With W_i being the solutions containing link l_i.

We have now arrived at the formulation for calculating link probabilities used in our method. In this discussion we have, for simplicity, not considered the uncertainty surrounding the

Platt scaling of the division probability estimates and the (dis)appearance probabilities, but including this does not influence the argument presented.

The approach detailed above relies on there being limited uncertainty regarding the values of our scaling factors. We justify this assumption by showing robust calibration across datasets and tight confidence intervals around the optimal value.

In the absence of a confident calibration, we can still assume our probability estimates are unbiased (a property of logistic regression), meaning that:

$$\mathbb{E}_{D_{GT}}(l_i|G_i, D_{GT}) = \mathbb{E}_{D_{GT}}(p_i)$$

This unbiasedness guarantees that if we use many different instances of our method calibrated on different ground truths to track different datasets and then combine the results the estimated error probabilities would be well-calibrated. The problem with uncertain calibration is therefore not so much that the estimates are 'wrong' as they are still maximally likely unbiased estimates. The problem is that that the quality of these estimates will correlate strongly between links. Where as in well-calibrated situations the probability of a link being true (l_i) given the prediction (p_i), $P(l_i|p_i)$, is independent from the truth value of other links, this is not true under calibration uncertainty. If the calibration is uncertain then finding a link to be true indicates that the others might be more likely to be true as well. This makes it difficult to treat individual datapoints obtained from the tracking (i.e a single cell cycle length) as truly independent and makes the derived biological conclusions uncertain, even if for a single link the estimated link probability remains unbiased.

In some situations, it might be hard to create ground truth datasets that allow for high confidence calibration. Specifically, when the marginalized predictions perform so good that the frequency of erroneous predictions is so low that only a few are present with respect to the ground truth data. In these cases, it might be advisable to take a reasonable higher bound for the scaling temperature (i.e the high side of the 95% confidence interval) such that the estimated error probabilities function as lower bounds.

3.6.2 Individual-wise bayesianity

Dietrich and List motivate multiplicative opinion pooling by showing it conforms to a strong intuition: that it should not matter if information is presented after the pooling procedure or beforehand to an individual predictor. They term this axiom 'individual-wise bayesianity' [105,106]. Following them we introduce a relative likelihood function $L(a)$, which encodes the extra information about an element a, such that updating our prediction after pooling would give:

$$P^L(a) = \frac{P(a)L(a)}{1 - P(a) + P(a)L(a)} \tag{1}$$

Following Dietrich & List, we can think of $L(a)$ in terms of conditional probabilities, i.e. the likelihood of some piece of new evidence (θ_a) occuring given element a being (un)true.

$$L(a) = \frac{P(\theta_a|a)}{P(\theta_a|\neg a)} \tag{2}$$

If instead the new information comes in the form of a new prediction, p_a, from a neural network or maybe a human annotator, then we can write the relative likelihood as:

$$L(a) = \frac{P(p_a|a)}{P(p_a|\neg a)} = \frac{P(a|p_a)}{P(\neg a|p_a)}\frac{P(\neg a)}{P(a)} = \frac{p_a}{1-p_a}\frac{1-P(a)}{P(a)} = \frac{L_a}{L_{a,prior}} \tag{3}$$

With L_a being the relative likelihood of element a being true given the predictions, as used throughout the main text ($L(p_a)$, in an alternative notation). $L_{a,prior}$ indicates the prior relative probability of element a based on the information the new prediction shares with already made predictions.

Assume now that this new information is known beforehand to the neural network (an individual predictor) assigned to predict event *a*. The consensus probability would then be given by:

$$P^L(a) = \frac{\sum_{\omega \ni a} L_{0,\omega}\, L(a) \prod_{i \in \omega} \frac{L_i}{L_{i,prior}}}{\sum_{\omega \setminus \ni a} L_{0,\omega} \prod_{i \in \omega} \frac{L_i}{L_{i,prior}} + \sum_{\omega \ni a} L_{0,\omega}\, L(a) \prod_{i \in \omega} \frac{L_i}{L_{i,prior}}} \tag{4}$$

$$P^L(a) = \frac{L(a) \sum_{\omega \ni a} L_{0,\omega} \prod_{i \in \omega} \frac{L_i}{L_{i,prior}}}{\sum_{\omega} L_{0,\omega} \prod_{i \in \omega} \frac{L_i}{L_{i,prior}} + (L(a)-1) \sum_{\omega \ni a} L_{0,\omega} \prod_{i \in \omega} \frac{L_i}{L_{i,prior}}} \tag{5}$$

If we then numerator and denominator by the partition function ($\sum_{\omega} L_{0,\omega} \prod_{i \in \omega} \frac{L_i}{L_{i,prior}}$), we get:

$$P^L(a) = \frac{L(a)P(a)}{1 + (L(a)-1)P(a)} \tag{6}$$

$$P^L(a) = \frac{P(a)L(a)}{1 - P(a) + P(a)L(a)} \tag{7}$$

Which is the same as equation (1) indicating that it does not matter at what stage information is introduced in our procedure and we conserve the 'individual-wise bayesianeity' axiom. It also provides us with a convenient way to incorporate new information about an element without redoing marginalization.

More generally we can see that integrating extra information about an element i during opinion pooling regarding any element a, is equivalent to updating the probability the following way after opinion pooling:

$$P^L(a) = \frac{P(a,i)L(i) + P(a,\neg i)}{1 - P(i) + P(i)L(i)} \tag{8}$$

This is in practice not a useful way of incorporating post hoc information as $P(a,i)$ and $P(a,\neg i)$ have to be computed using our marginalization approach as well.

In our derivations above we assume that the extra information shares zero overlap with the previously available information. If this would not hold we should discount $L(a)$ by dividing it by an appropriate temperature as proposed in the main text and method section, but this would not affect our argument.

3.6.3 Bayesian belief matrix

Another way of integrating multiple probabilities is using a Bayesian belief matrix approach [114]. Here information is integrated sequentially using Bayesian update rules. Per integrated prediction the process has two steps. First we update the belief about the element the prediction is made on:

$$P_{n+1}(a) = \frac{P_n(a)L(a)}{1 - P_n(a) + P_n(a)L(a)} \tag{9}$$

We then normalize all other probabilities to satisfy the constraints of the tracking problem. This method is limited in which type of constraints it can take into account, because all constraints have to be implemented in a single normalization step. It can therefore only deal with mutually exclusive events, such a links connecting towards same node under the condition that cell merging is prohibited. For every event i mutually exclusive with element a, the updated probabilities become:

$$P_{n+1}(i) = P_{n+1}(i)\frac{1 - P_{n+1}(a)}{1 - P_n(a)} \tag{10}$$

If we then substitute $P_{n+1}(a)$:

$$P_{n+1}(i) = P_{n+1}(i)\frac{1 - \dfrac{P_n(a)L(a)}{1 - P_n(a) + P_n(a)L(a)}}{1 - P_n(a)} \tag{11}$$

$$P_{n+1}(i) = P_{n+1}(i)\frac{\dfrac{1 - P_n(a)}{1 - P_n(a) + P_n(a)L(a)}}{1 - P_n(a)} \tag{12}$$

$$P_{n+1}(i) = \frac{P_{n+1}(i)}{1 - P_n(a) + P_n(a)L(a)} \tag{13}$$

This is identical to equation (8) for mutually exclusive events i and a, in which case $P(a, i)$ is zero and $P(a, \neg i)$ is the same as $P(a)$. We can thus see that the 'Bayesian belief matrix' update rules are simply the same as the rules for integrating information post-hoc in our multiplicative pooling framework. 'Individual-wise bayesianeity' then guarantees, that the stepwise updating via Bayesian update rules is mathematically equivalent to our marginalization approach in the case of mutually exclusive events. The starting probabilities $P_0(i)$ function in this case as the priors.

Our approach can even be called a generalization in the sense that the Bayesian updates can only deal with elements that are mutually exclusive, otherwise the normalization step would not work. While our method can also deal with more advanced constraints, such as that two links can exit a node but only if the cell is dividing.

Note that in practice our method and the one described by Ulicna et al. function quite differently. In their method there are no neural networks estimating probabilities of events from the observed data. Instead they work with models that give the probability of observing data given the event we want to predict. Their version of equation (10) would thus take the form of:

$$P_{n+1}(a) = \frac{P_n(a)\frac{P(\theta_a|a)}{P(\theta_a|\neg a)}}{\big(1 - P_n(a)\big) + P_n(a)\frac{P(\theta_a|a)}{P(\theta_a|\neg a)}} \tag{8}$$

The probability of observing the data when an event (like a link) is not true $P(\theta_a|\neg a)$ in these type of methods is often intractable and they not used directly. Instead, this probability can be assumed to be equal for all events in the system. The fact that all events are mutually exclusive then allows $P(\theta_a|\neg a)$ to be simply multiplied out, as multiplying all the relative likelihoods by the same constant does not affect the outcome, making the update rule:

$$P_{n+1}(a) = \frac{P_n(a)P(\theta_a|a)}{\big(1 - P_n(a)\big) + P_n(a)P(\theta_a|a)} \tag{10}$$

This assumption of equal $P(\theta_a|\neg a)$ is reasonable for similar events (different links), but clearly not when considering for instance the chance of a cell disappearing $P(d)$. This problem is ameliorated by Ulicna et al. by capturing it in a single hyper parameter, $P^*(\theta|d)$, by which to divide all link likelihoods:

$$P^*(\theta|d) = \frac{P^*(\theta|d)P(\theta_a|\neg a)}{P^*(\theta|\neg d)} \tag{11}$$

This factor integrates the chances of cell disappearances in the predictions, without needing to explicitly update them. This then gives the following update rule:

$$P_{n+1}(a) = \frac{P_n(a)P(\theta_a|a)}{\big(1 - P_n(a)\big)P^*(\theta|d) + P(a)P(\theta_a|a)} \tag{12}$$

The fact that $P(\theta_a|\neg a)$ is not used directly reflects the fact that it is difficult to define models for these types of probabilities. This highlights an advantage of neural networks that instead aim to estimate the probability of an event given the data $P(a|\theta_a)$. The likelihood information then simply becomes, like in equation (3)):

$$L(a) = \frac{P(\theta_a|a)}{P(\theta_a|\neg a)} = \frac{P(a|\theta_a)P(\neg a)}{P(\neg a|\theta_a)P(a)} = L_a(\theta_a)/L_{a,prior}$$

In which an uniform prior could be chosen like in Ulicna et al. or it can be chosen to reflect possible shared information as described in our methods section. The final outcome of the 'Bayesian belief matrix' approach can then simply be calculated as (with $L(d)$ giving the relative likelihood of a cell disappearing):

$$P_{final}(a) = \frac{L(a)}{\sum_i L(i) + L(d)} \tag{13}$$

4 Organoid cell fate dynamics in space and time

Organoids are a major new tool to study tissue renewal. However, characterizing the underlying differentiation dynamics remains challenging. Here, we developed TypeTracker, which identifies cell fates by AI-enabled cell tracking and propagating endpoint fates back along the branched lineage trees. Cells that ultimately migrate to the villus commit to their new type early, when still deep inside the crypt, with important consequences: 1) Secretory cells commit before terminal division, with secretory fates emerging symmetrically in sister cells. 2) Different secretory types descend from distinct stem cell lineages rather than an omni-potent secretory progenitor. 3) The ratio between secretory and absorptive cells is strongly affected by proliferation after commitment. 4) Spatial patterning occurs after commitment through type-dependent cell rearrangements. This 'commit-then-sort' model contrasts with the conventional conveyor belt picture, where cells differentiate by moving up the crypt-villus axis, and hence raises new questions about the underlying commitment and sorting mechanisms.[1]

[1] The content of this chapter is published as: X. Zhang, M.A. Betjes, P.E. Ender, Y.J. Goos, G. Huelsz-Pince, H. Clevers, J.S. van Zon, S.J. Tans, *Science Advances* (2023).

4.1 Introduction

The past decade has brought groundbreaking advances in organoid technology, which enabled many new approaches to studying organ development, self-renewal, and pathology[38,135-138]. The benefit of organoids lies in their ability to *in vitro* generate the diverse cell types and 3D organization that characterize *in vivo* organ tissue. Key features of the cellular architecture have been established using immunostaining, fluorescence microscopy, and single-cell RNA sequencing[17,46,70,139,140]. However, these methods do not reveal the dynamics that are critical to the underlying cellular reorganization and differentiation – as exemplified for instance by the rapidly renewing intestine. Lineage tracing[141-144] can visualize the offspring of some cells but not when or where cells adopt a new fate, nor their movements or lineage trees. In parallel, important advances have been realized in 3D time-lapse microscopy. Confocal and light-sheet imaging can visualize the dynamics of cell proliferation and migration[39,44,46,145,146], while machine learning methods can track cells in time[44,53,89,91]. Yet, this approach has been of limited use in studying the dynamics of differentiation. While fluorescent protein markers can in principle be used to identify cell types[61,147,148], the many involved types (over 5 in intestinal organoids)[17] and the fluorescent labeling of nuclei required for tracking present major challenges in terms of phototoxicity, genetic engineering, and wavelength overlap.

Hence, the dynamic basis of cellular organization has not been resolved for any organoid system[85]. In intestinal organoids, stem cells are thought to differentiate when moving into the transit amplifying (TA) region higher up in the crypt, driven by proliferation at the crypt bottom, upon which they stop proliferating and continue upwards to the villus region, where most differentiated cells reside and are ultimately shed into the lumen[86,149-151]. Cues are provided by WNT activity that decreases along the crypt-villus axis and Notch inhibition between secretory and absorptive types that regulate their relative abundance[152-154]. This 'conveyor belt' model is appealing because it can naturally explain how different types are dynamically organized in space: cells driven towards the differentiated villus region are triggered to differentiate, while cells remaining in the crypt preserve their stem identity. However, with dynamic data lacking, it remains unclear if this principle holds and whether other mechanisms contribute. Addressing these issues requires the characterization of all cells within key regions.

We, therefore, developed a method that integrates type identification and tracking (TypeTracker). It works by propagating cell types back in time along lineage trees, using machine learning-based analysis of 3D video microscopy, multiplexed antibody staining of endpoint cells, and a rule-based backpropagation protocol (**figure 4.1A**). The resulting multi-dimensional data identifies the moment and location where cells differentiate, their movements and genealogical relations to other cells, and the neighbors they interacted with. It can be used for any organoid with proliferating and differentiating cells and requires only confocal microscopy, immunostaining, and data integration software.

Application to mouse intestinal organoids[135] (**figure 4.1B)** revealed an unexpected principle for the interplay between differentiation and spatial organization. Type commitment occurred notably early: well before the differentiating cell lineages ceased to proliferate and moved upward to the villus region – at a moment when the cells still occupied the same region as their stem-cell relatives that continued to retain the stem identity and remained deep inside the crypt. This early commitment had several notable features. Secretory fate emerged typically with identical types in pairs of sister cells, indicating their mothers had committed already, and contrasting with the notion that committed cells arrest their cell cycle[155-157]. While cells of different secretory types were observed, they were not closely related and rather emerged from distinct stem cell lineages. This finding challenges the existence of omnipotent secretory precursors that produce progeny of different secretory types[155,156,158-161]. Absorptive lineages are committed even longer before cell cycle arrest (over 5 divisions), providing an explanation for the high abundance of absorptive over secretory types. In contrast with current models[86,162,163], the stereotypical positioning of cell types was achieved through spatial segregation after commitment, and secretory sister cells separating more rapidly than neighboring cells of different types. Given the conserved nature of the underlying pathways such as WNT and Notch, this 'commit-then-sort' mechanism may well be widely relevant. Our TypeTracker approach can be used broadly to elucidate the dynamic differentiation programs of organoid systems, and how they are affected by mutations or external conditions and can be integrated with live cell monitoring of key cell biological processes.

4.2 Results

4.2.1 Overview of TypeTracker.

Our method consists of four main steps (**figure 4.1A**). First, movies of organoids growing in Basement Membrane Extract (BME) gel are recorded by 3D confocal microscopy (**supplementary figure 4.1A-C**). We focus on intestinal organoids and image complete organoids or their crypt protrusions every 12 minutes for 60 hours, using an H2B-mCherry marker to detect cell nuclei (**supplementary figure 4.1D**). Second, we employ a convolutional neural network[53] to track the nuclei in space and time across several generations (**figure 4.1B, supplementary figure 4.1E, F**), and visualize their genealogical relations using lineage trees (**figure 4.1C, supplementary figure 4.2, supplementary figure 4.3**). The tracked organoids had different sizes, but all showed similar cell cycle durations. Consistently, proliferative lineages, which display up to seven divisions, are mostly found at the crypt bottom, while non-dividing cells are abundant in the villus-like region of the organoid. Third, we perform multiple rounds of antibody staining to determine the cell types at the movie endpoint, as detailed below **(figure 4.1D, E)**. Hence, we identify the main types found in the intestine: stem, enteroendocrine, Paneth, goblet, enterocyte, transit-amplifying,

as well as an immature mucus-producing type. Fourth, we infer the cell types at earlier time points, starting at the endpoint of the lineage tree and progressively moving back along each branch to the beginning, using a set of rules that we describe in detail below. Hence, one obtains virtual organoids that present the spatial location, division events, movements, and type of all cells (**figure 4.1F**). The completeness of this method, which identifies all cells in a region of interest and associated lineage trees, enables our backpropagation approach and analysis of spatio-temporal correlations.

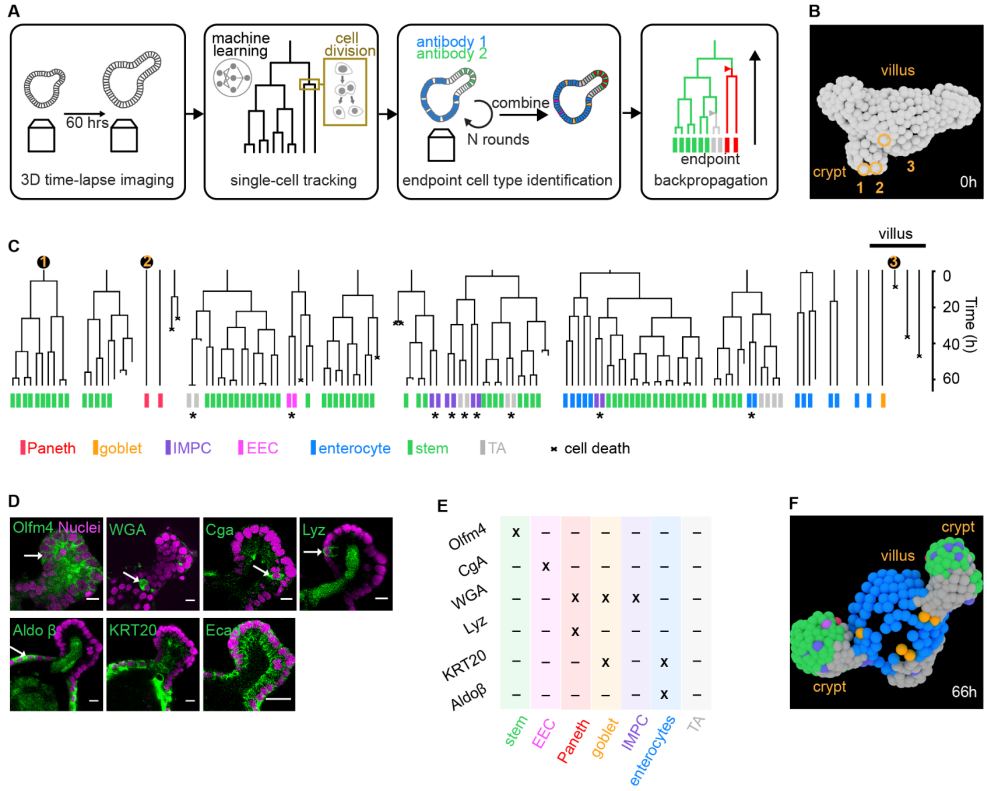


Figure 4.1: TypeTracker method overview and primary data. **A,** Method workflow. Step 1: Image several organoids in 3D for 60h. Step 2: Track all cells tree using convolutional neural network. Each cell is uniquely identified from frame to frame, yielding a lineage (black line). Division events give a fork-like structure. Step 3: Identify cell type at movie endpoint using multiple rounds of antibody staining. Step 4: Map cell types onto lineage tree and infer historical cell types including type transitions (triangles). **B,** Reconstructed cell positions at start of lineage tree. **C,** Example lineage trees and their mapped endpoint cell types. Numbers: lineages of three cells seen in panel B. Stars: endpoint sister pairs of same type, but of different type than cousins, suggesting type transition in mother cell. **D,** Nuclear marker (magenta) and antibody or dye staining (green): Olfactomedin 4 (Olfm4), Wheat Germ Agglutinin (WGA), Chromogranin A (Cga), Lysozyme (Lyz), Aldolase β (Aldoβ), Cytokeratin 20 (Krt20) and Ecadherin (Ecad). **E,** Cell type identification table. **F,** Reconstructed cell positions and type at end of the lineage tree (panel C), 66 hours after the start (panel B).

4.2.2 Cell type identification.

We developed an approach to map cell type data, as obtained by several rounds of antibody staining, onto the endpoint of the movie and lineage trees (**figure 1D, E**). This method includes washing protocols that minimize organoid deformations, antibody stripping between rounds, and a min-cost flow solver algorithm[99] combined with manual correction to link cells in the staining images to the cells at the movie endpoint. Stem cells were identified by Olfactomedin 4 (Olfm4)[164], enterocytes by Aldolase β (Aldoβ)[46] and enteroendocrine cells (EECs) by Chromogranin A (Cga)[165]. Consistently, Olfm4+ cells were found only at the crypt bottom, Cga+ cells in both the crypt and villus regions and Aldoβ+ cells exclusively in the villus region (**figure 1F**). Stem cells and EECs did not stain for any of the other markers used, while enterocytes also stained for Cytokeratin 20 (KRT20)[166], as expected. Wheat Germ Agglutinin (WGA), which stains intestinal mucus[167,168], labeled cells in the crypt and villus regions. A subset of these were also positive for Lysozyme (Lyz) and thus identified as Paneth cells[169], while another subset expressed KRT20 and was identified as goblet cells[166]. Consistently, the former often contained granules that are typical of Paneth cells, while the latter had a cup-shaped morphology that characterizes goblet cells (**supplementary figure 4.4**). A remaining subset of WGA+ cells expressed neither Lyz nor KRT20. Evidence provided below suggests that these are immature Paneth and goblet cells, which we thus refer to as immature mucus producing cells (IMPCs). We identified a population of cells within the crypt that did not stain for any of our markers, which are identified as transit-amplifying (TA) cells based on evidence discussed below.

4.2.3 Sister cells adopt the same type.

Prominent in our data is that endpoint types generally came in pairs, with sister cells consistently displaying the same fate (**figure 4.1C**). Statistical analysis showed that 97% of all endpoint sister (N=869 sister pairs in 9 organoids) were of the same type. All types, including secretory types, exhibited this symmetry, as evidenced by the dominant diagonal in the sister-sister cell type histogram (**figure 4.2A**). This consistent type-symmetry between sisters is notable. Intestinal stem cells are proposed to exit from their cell cycle when committing to a secretory fate[155,156], which would rather yield asymmetry, as one sister may then adopt a secretory fate that the other does not. External cues can, in principle, produce a local environment that is similar for the two sister cells, and hence drive them to the same fate. However, a priori there is then no reason why only sisters should be impacted, as also non-sister cells like cousins can be neighbors. Importantly, the fates of many sister pairs were different than their direct cousins (**figure 4.1C**, stars). This showed that the underlying type transitions did affect sisters specifically and occurred during the observed growth period of the lineage. These data indicated that the type transitions rather occurred in the mother (or earlier generations), which subsequently divided to produce two daughters of the same type. Larger subtrees, which for instance showed two or more sister pairs of the same type (**figure**

4.1C, blue enterocyte subtree, **supplementary figure 4.2, supplementary figure 4.3**), were consistent with divisions occurring after type commitment and hence producing cells of the same type.We similarly observed symmetry between goblet sisters in conditions that enriched for goblet cells (IWP2 and DAPT)[147,170]. For those goblet lineages that showed a division, the sister cell also adopted the goblet fate (**supplementary figure 4.5**). We also obtained more evidence for symmetry between Paneth sister cells (see below) and Enteroendocrine (EEC) cells: a fluorescent reporter of Neurogenin 3 that marks EEC cells[147], was found to increase simultaneously in two nearby cells (**supplementary figure 4.6**).

4.2.4 Dependence of cell age on differentiation state.

The age of the cells at the movie endpoint was defined as the time since the last observed division (**figure 4.2B**). Notably, the TA cells showed a statistically identical age distribution as the stem cells, in line with the notion that they are highly proliferative[2,156,162]. EECs, enterocytes, Paneth, and goblet cells also showed a broad distribution, but with an average age that was substantially larger than the stem and TA cells, consistent with their terminally differentiated nature. The ages of these terminally differentiated cells were rarely below 20 hours (**figure 4.2B**), suggesting that the involved markers (Cga, AldoB, Lyz and KRT20) took a substantial amount of time to become observable.

4.2.5 Cell type backpropagation.

To infer the cell type along the lineages of the family tree, we used the lineage tree topology and the above observations (**figure 4.2A, B**). Starting at the lineage endpoints, we propagated the endpoint types back in time using this process: 1) From one timepoint to the previous timepoint, the type is initially assigned as unchanged if branch points (divisions) are not traversed but may later be updated when a cell type transition is identified. 2) At branch points, the two involved daughters are considered: if both display the same (inferred) type, the mother is assigned as that type, as discussed above (**figure 4.2A**). 3) If one daughter is a stem cell and the other is not, the mother is assigned as a stem cell, based on the notion that stem cells are generated from stem cells. 4) If one daughter is a TA cell and the other is not, nor a stem cell, the mother is assigned as a TA cell. This rule is based on the TA cells being proliferative like stem cells (**figure 4.2B**), which makes the alternative (the mother has the identity of the non-TA daughter) unlikely. In the last two cases (3 and 4), a type-transition is inferred in one of the daughters. We mark the latter with an arrow halfway the cell cycle (**figure 4.2C**). These rules allowed the backpropagation of complete trees from the end to the beginning (**figure 4.2D, supplementary figure 4.7**).

Next, we estimated stemness by quantifying the fluorescence intensity of Lgr5-GFP, a well-known stem-cell marker[171], to further test our method. Phototoxicity, which is stronger for GFP than for mCherry, limits the imaging frequency and duration. GFP at the movie endpoint indeed correlated with measured Olfm4 staining intensities (**supplementary figure 4.8A, B**).

Backpropagation-inferred stem cells indeed showed higher GFP expression, while inferred enterocytes or TA and goblet cells showed lower GFP (**supplementary figure 4.8C, D**). Notably, lineages inferred to go from stem to TA type showed decreasing GFP, and those decreases were established in the mother cell – identified as the committing cell by the TypeTracker method (**supplementary figure 4.8E**), thus supporting the backpropagation approach.

Paneth cells also showed consistency between backpropagated type and real-time visualization, by identifying the granules that characterize Paneth cell (**supplementary figure 4.9A**). The latter also showed Paneth cell division (**supplementary figure 4.9A**). Next, we performed time-lapse imaging with a Paneth cell marker (CD24) (**supplementary figure 4.9B)**. Most CD24+ cells did not show divisions, as expected for terminally differentiated Paneth cells. We did see cells that divided and then became CD24+. Notably, the corresponding sister cell then also showed the CD24 signal increasing in the same manner (**figure 4.2F, supplementary figure 4.9C, D**). These data are consistent with the mother committing to the Paneth fate and CD24 expression following later in the daughters. These findings highlight the advantage of our method when markers are visible only long after commitment and hence not informing on early differentiation events. Note that limitations of our method are described in the discussion.

4.2.6 Differentiation pathways.

The observed differentiation pathways indicated notable features (**figure 4.2D, supplementary figure 4.7**). Trees typically displayed 1 to 3 cell type transitions, thus yielding sub-trees of the new type, while the old type was maintained in another sub-tree, and some lineages showed two consecutive type transitions. For instance, a stem cell tree was first shown to spawn a TA-subtree, which in turn generated an enterocyte subtree (**figure 4.2D**, tree 1). This order, in which TA is an intermediate type between stem and terminally differentiated types[86], is consistent with the sister-sister cell type histogram: besides the dominant diagonal, low-frequency off-diagonal entries indicated stem-TA sister pairs, and a few pairs of one TA and one terminally differentiated type, but never sister pairs showing two different terminally differentiated types (**figure 4.2A**). The latter would be in line with models where stem cells first differentiate into secretory precursors, which in turn can generate different terminally differentiated types[155,156,158-161]. To further probe this notion, we extracted the largest possible subtrees from our data that contained two types, and found that none had two different secretory types, nor one secretory and one absorptive type. Instead, they typically combined TA and either stem (52.2%) or a terminally differentiated state (31.4%) (**supplementary figure 4.10**), consistent with the sister relations (**figure 4.2A**). Among these subtrees, we also found cases (16.4%) that combined stem and terminally differentiated states (**figure 4.2D**, tree 2 and 3), suggesting that these lineages have a negligible TA role.

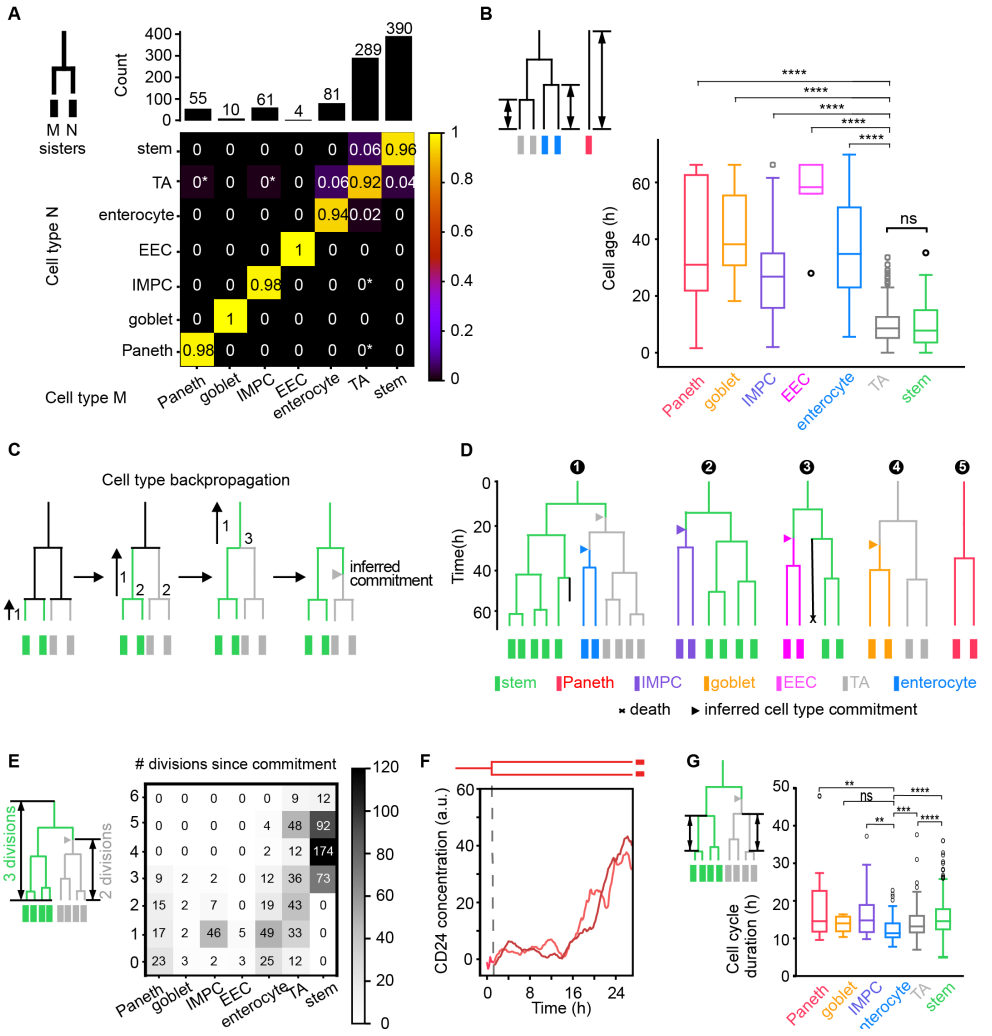


Figure 4.2: Differentiation pathways and cell type commitment. **A,** 2D histogram indicating the two cell types of sisters at the movie endpoint, as determined using antibody staining. Shown is the number of times a cell type combination was observed divided by the total of each column, indicated on top. Star: low frequency resulting from a single observation. The dominant diagonal shows sisters typically have the same type. **B,** Age of stained cells, quantified as the time since the last division, or since movie start for lineages without division. ****, $p \leq 0.0001$; ns, $p > 0.05$. TA cells show the same age as stem cells, indicating their proliferative nature. Consistently, the differentiated cells (EECs, enterocytes, Paneth and goblet cells) are older. **C,** Cell-type backpropagation rules. 1: working backwards along each lineage, the inferred type is unchanged when not encountering branchpoints (divisions). 2: for branchpoints, if both daughters are the same type, the mother is assigned that type. 3: If only one daughter is a stem cell, the mother is assigned as a stem cell. 4 (not shown): If only one daughter is a TA cell, and the other is not a stem cell, the mother is assigned as a TA cell. In cases 3 and 4, a type-transition is inferred in one daughter (triangle). **D,** Example lineage trees as determined by TypeTracker. **E,** Number of observed divisions for each cell type. **F,** Cell division generating two Paneth daughters showing similar profiles of increasing CD24 signals. **G,** Cell cycle duration for different cell types. Cell cycle durations were notably similar between types, while enterocytes exhibited lower mean cell cycle durations than TA cells, stem cells and the secretory cell types. All data in this figure were from nine organoids.

4.2.7 Division rather than differentiation rates control enterocyte abundance.

Cell type transitions to enterocytes were less frequent than to the secretory types combined (about 1.5-fold), even as enterocytes outnumbered secretory cells. To investigate this issue, we quantified the number of consecutive divisions in each differentiation state (**figure 4.2E**). The EEC, goblet and IMPC secretory states mostly showed one division, sometimes two. In contrast with common thinking[155,156], we found Paneth cells to divide as well (**figure 4.2E**). The inferred Paneth cell divisions were indeed confirmed by the continuous presence of granules that are specific to Paneth cells and are observable without labeling (**supplementary figure 4.9A**). Next, we quantified the cell cycle duration for each cell type. Notably, these complete cell cycles were similarly long for all types including secretory and stem types (**figure 4.2G**). The lineages were observed to stop dividing subsequently, consistent with a cell cycle exit.

To further probe proliferation of secretory cells, we co-stained with the Paneth marker Lysozyme and the proliferation marker Ki67. Consistently, 33% of the Lysozyme+ cells were indeed also Ki67+ (**supplementary figure 4.11A**). We also studied cells with a Neurogenin 3 label that marks EEC cells, and found that a small fraction was Ki67+, consistent with committed EEC secretory cells capable of dividing (**supplementary figure 4.11C**).

Enterocyte lineages displayed significantly more divisions than secretory cells. The former reached to up to 5 divisions, not far off from the 6 divisions seen for stem and TA states (**figure 4.2E**). Hence, after specification, absorptive lineages generated substantially more cells than secretory lineages, which thus offset their comparatively low transition rate. This finding is consistent with recent data in which absorptive lineages showed larger clone sizes than secretory lineages[143,144]. Enterocytes were even found to exhibit the lowest mean cell cycle duration (12.5 h), significantly lower than for both stem (15.5 h) and TA cells (14.1 h) (**figure 4.2G**). Overall, these results indicate that relative abundance is controlled by the cell lineage dynamics, in particular the number of divisions after specification.

4.2.8 The spatio-temporal differentiation program.

Access to spatial dynamics is a key benefit of our approach. The positions of inferred cell types mapped along the crypt-villus axis as expected, with stem and Paneth most towards the bottom, followed by IMPC, EEC, TA, enterocyte, and goblet[86,150] (**figure 4.3A, B, supplementary figure 4.12A**). In addition, we can now map the point of commitment onto the spatial organoid structure. Commitments to secretory types were broadly distributed along the crypt-villus axis, but typically occurred deeper in the crypt than absorptive commitments (**figure 4.3C**). The latter also distributed broadly and were positioned well within the crypt. Notably, when we mapped full lineage trees along the crypt-villus axis, we found that these committing cells were still in the same approximate spatial region as their close relatives that maintained the stem fate (**figure 4.3D, E, supplementary figure 4.12B,**

C). Spatial separation from these stem-cell relatives did occur after committing (**figure 4.3D, E**). This separation was most evident for lineages destined for the villus region (enterocytes and goblet cells), with the stem lineages they arose from remaining confined to the crypt (**figure 4.3D, E**, top trees). Such separation from relatives was not always observed, specifically for cells committing to a secretory type (IMPC) near the crypt bottom and remaining there (**figure 4.3D** bottom tree, **figure 4.3E**, top tree). Consistently, the mean migration speed is highest for enterocytes and goblet cells, and lowest for secretory fates like Paneth cells and IMPC's (**figure 4.3F**)

4.2.9 Promotion of sister separation.

Secretory types were typically born as sisters (**figure 4.2A**) and, hence, as neighbors. We wondered how this finding relates to the observation of single Paneth being surrounded by stem cells[135,172]. In our data, most neighbors that were *not* sisters (and hence can be of different type), were still neighbors after 2h (85%) and 10h (57%), independently of cell type (**figure 4.3G, H, supplementary figure 4.12D**). Neighbors that *were* sisters, which almost always were of the same type, separated more frequently after 2h (71% remains neighbors), likely due to arrangements directly following divisions[45,173]. Surprisingly, Paneth and IMPC sisters showed even stronger separation over longer timescales, and more than the other types, with less than 20% still neighbors after 10h (**figure 4.3I**). Sisters, and in particular of these secretory types, thus appear to rearrange more strongly to achieve interspersion. These findings show that active cell separation after fate commitment contributes to cell type patterning in the epithelium.

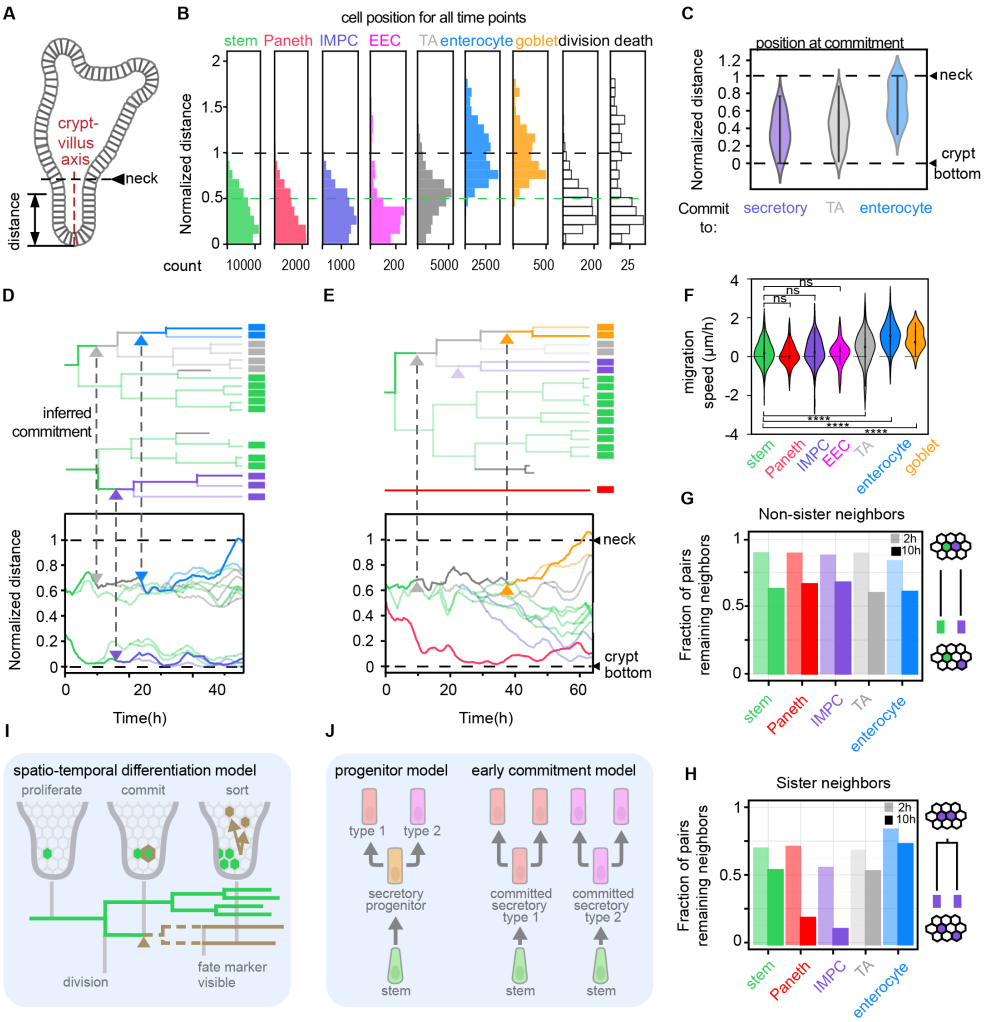


Figure 4.3: Spatio-temporal differentiation program. A, Organoid schematic diagram. **B,** Inferred cell types mapped along the normalized crypt-villus axis. **C,** Position of cells when they commit to a new type. **D & E,** Lineage trees mapped in space along the crypt-villus axis. Bottom trees: Transition to IMPC and position of Paneth cell deep in the crypt surrounded by stem cells. Top trees: Transitions to enterocytes or goblet cells higher in the crypt, but close to their stem cell relatives, followed by their movement to the villus region, rather than the other way around. **F,** Cell migration speed, as quantified by the migration distance along the crypt-villus axis after commitment or movie start, divided by the migration time. **G,** Cell separation, for one cell of the specified type and one cell of any type. Displayed is the fraction of pairs that remain neighbors in 2 or 10 h. Most neighbors remained neighbors even after 10h, independently of cell type. **H,** Cell separation of sisters of the same type, Sisters are shown to separate faster than non-sisters (panel g), particularly sister Paneth cells and IMPCs. **I,** Commit-then-sort model. Cells commit early (to unspecified fate, brown triangle), when still nearby their stem cell relatives, and after which they can still undergo division. Tested fate markers rise in detectable manner after a certain time only. Spatial patterning occurs after commitment by type-dependent rearrangements. **J,** Differentiation pathway models. In the first, stem cells first become an omni-potent secretory progenitor that can divide and produce different secretory types. In the second, cells commit early to different secretory types and produce cells of the same type when they divide. Our dynamic data is consistent with the latter model.

4.3 Discussion

Before discussing the findings, we review the limitations of our approach. First, transient fates like stem and TA present early in the tree may be missed, as their type may not be maintained until the endpoint (**supplementary figure 4.13A, B**). This issue is exacerbated by cell death events (crosses), as aborted lineages cannot propagate type information back in time (**supplementary figure 4.13A, C**), and partly mitigated by the fact that stem and TA types are proliferative and hence generate multiple branches. Second, if cells would show reversible transitions back to the stem or TA type, not all these events may be detected, again owing to their transient nature. Third, used antibodies may not identify immature types, which have committed briefly before the end of the movie. This issue impacts marker-based methods more generally. Our data indeed suggests lengthy type maturation, as evidenced by the dominant sister symmetry (**figure 4.2A**), high age of differentiated cells at movie endpoint (**figure 4.2B**), and live-cell marker data (**figure 4.2F, S6**). Our method offers advantages in overcoming this limitation, by detecting early commitment using lineage relations even when the phenotype is not detectably expressed yet.

Our data showed many features that are consistent with existing models of intestinal organoid renewal, including the presence of stem cell and TA zones, expected positions of differentiated cell types, and an effective migration along the crypt-villus axis driven by proliferation in the crypt bottom[86,150,163] (**figure 4.3A, D-F**). The temporal order of events that our dynamic observation addresses was notable, however. Specifically, we found cells committing to a new type when still positioned in the same area as their stem cell relatives, well before their separation into the cell lineages that move towards the villus region and those that remain in the crypt (**figure 4.3C-E, 3I**). Indeed, the opposite was not observed: cells that first move away from their stem cell relatives and then differentiate when they have moved upwards to the villus region, as in the classic conveyor belt model. We found commitment even deeper in the crypt for secretory types such as Paneth cells that ended up at the crypt bottom, thus not requiring the proposed downward migration after commitment[86,174]. This observation agrees with a recent mouse study reporting that both Paneth cells and enteroendocrine cells arose directly from stem cells at the +4/+5 positions in the intestinal crypt[173]. Another striking feature was the rapid separation of neighboring secretory sisters (**figure 4.3G-I**), showing how local cell rearrangement can contribute to spatial cell type patterning. Such rearrangements may explain observed non-neighboring positions of clonal cells in the mouse intestine[175]. Overall, our data indicated a picture in which cells first commit and subsequently spatially sort and reorganize (**figure 4.3I**).

This early commitment picture has other consequences. One example is the highly symmetric adoption of identical fates by sister cells, including those with secretory fates (**figure 4.2A**). It has been proposed that secretory differentiation occurs through an omni-potent secretory precursor that can generate all types of secretory cells[155,156,158-161] (**figure 4.3J**). Our early commitment findings suggest that this omni-potent secretory state, if it exists, occurs only

briefly and is exclusively restricted to the mother cell, as different secretory fates are not seen among the mother's offspring (**figure 4.3J**). This conclusion was supported by the more distant relatives: subtrees containing two types at most, never contained two different secretory types (**supplementary figure 4.10**). An omni-potent secretory state was also questioned by a recent quantification of gene expression versus pseudo-time, which indicated differentiation via lineage-specific unipotent transitions[173].

Our results indicate that type commitment occurs in mother cells that still undergo one or more divisions and transmit this type to its descendants (**figure 4.3J**). These findings contrast with the idea that cells exit their cell cycle upon committing to a secretory fate[155-157], and suggests division may be important to completing differentiation[176]. Live cell secretory markers (CD24 and Neurogenin 3) were not visibly expressed in the mother, and rather rose simultaneously in the two daughters, in line with a committing mother (**supplementary figure 4.6, supplementary figure 4.9**). These data may also help explain why previous work suggested that committed secretory cells are non-proliferative[155-157]. Reported clonal expansion of goblet cells after fate commitment in mouse colon is consistent with our findings[143]. Notch signaling between neighbors is proposed to restrict secretory fate to a single cell that is surrounded by non-secretory cells[177]. Our observation of neighboring sisters with the same secretory fate (**figure 4.3H-I**) implies that Notch signaling affects cell fate early and transiently, likely in the mothers of secretory cells.

We note several differences between organoid and *in vivo* development. Organoid crypts and villi lack the supporting mesenchymal tissue and hence the associated space constraints and range of molecular signals. *In vivo*, the mesenchyme can maintain stem cells in absence of the Paneth cells[178-182], suggesting a redundancy in Wnt signaling that can affect aspects such as the size of the stem cell zone. In addition, organoids have an enclosed lumen that accumulates debris over time, which limits the observation time and could affect cellular metabolism. The ratio of differentiated and stem cells the ratio of absorptive and secretory cells may also not directly compare between organoids and *in vivo*[183], though our data can show how increased divisions after enterocyte specification can be exploited to control their relative abundance over secretory cells (**figure 4.2E**). Finally, organoids lack many other factors, including pathogens and interacting immune cells, which can promote the rare tuft and M-fold cells.

Our method can be applied broadly to study the spatio-temporal differentiation programs of organoid systems, and how they are impacted by external conditions such as metabolic compounds or interleukins, interacting cell types including bacteria and immune cells, and disease mutations. Its focus on spatial and temporal characterization makes it distinct from and complementary to other methods such as single cell RNA sequencing[17,70] or multiplexed tissue imaging[184,185]. The TypeTracker approach follows systematically all cells in a region of interest, thus allowing direct correlative analysis, and is straightforward to implement as it requires only confocal microscopy, antibody staining, and the algorithms we present here.

It can also readily be combined with other measurements, such as endpoint Single Molecule Fluorescence In Situ Hybridization (smFISH) and real-time fluorescence microscopy of various cellular processes and signals.

4.4 Materials and Methods

Organoid culture. Murine intestinal organoids carrying both a H2B-mCherry reporter and a Lgr5-GFP reporter were gifts from Norman Sachs and Joep Beumer, from the group of Hans Clevers in Hubrecht Institute. Organoids were embedded in 'domes' of basement membrane extract (BME, Trevingen) in tissue culture plates. They were further submerged in growth medium consisting of murine recombinant epidermal growth factor (EGF 50 ng/ml, Life Technologies), murine recombinant Noggin (100 ng/ml, Peprotech), human recombinant R-spondin 1 (500 ng/ml, Peprotech), n-Acetylcysteine (1 mM, Sigma-Aldrich), N2 supplement (1x, Life Technologies) and B27 supplement (1x, Life Technologies), Glutamax (2 mM, Life Technologies), HEPES (10 mM, Life Technologies), Penicilin/Streptomycin (100 U/ml 100 µg/ml, Life Technologies) in Advanced DMEM/F-12 (Life Technologies). Organoids were kept in incubators at 37 °C and with 5 % CO_2. The medium was changed every two days. Each week, organoids were passaged by mechanically dissociating crypts using a narrowed glass pipette.

Organoid sample preparation for imaging. In conventional culture conditions, organoids were embedded in 'domes' of BME droplets and were thus at different heights relative to the plate bottom (**supplementary figure 4.1A**). Imaging organoids located far from the plate bottom required long working distance objectives and increased light exposure, leading to excessive phototoxicity. To improve the imaging procedures, we used the 4 well chambered cover glass (# 1.5 high performance cover glass) from Cellvis as imaging plates. Organoids were broken into single crypts, seeded in the imaging plates and put in fridge (~ 4 °C) for ~ 10 minutes, allowing them to sink downwards to the cover glass. Afterwards, they would be incubated at 37 °C with 5 % CO_2 for 20 minutes so that the gel could solidify with organoids settled at the bottom of the wells (**supplementary figure 4.1B**). Growth medium was added after the incubation. Organoids were then kept in the incubator for around 2 days until the imaging experiments. For the enrichment of goblet cells, 5uM IWP2 and 10uM DAPT were added to the medium.

Time-lapse imaging with 3D confocal microscope. Time-lapse imaging was performed with a scanning confocal microscope (Leica TCS SP8) with a 40x water immersion objective (NA = 1.10). Experiments were performed at 37 °C and 5 % CO_2. More than 20 organoids with already budded crypts were selected for imaging. Stacks of ~ 30 z-slices with 2 µm step size were taken every 12 minutes per organoid. Imaging of H2B-mCherry was conducted with an excitation laser of 552 nm at 1 % of the laser power and the emission signals were collected with Leica HyD hybrid detectors whose filter range was set to be 557 nm - 789 nm.

Live-cell tracking. Live-cell tracking was conducted by OrganoidTracker, a software developed by our group[53]. The positions of each nucleus were predicted with a trained neural network and cells were then automatically linked between frames based on the relative

positions and nuclear sizes. The software could report warnings when the linking was less reliable and allowed for manual corrections.

3D reconstruction. 3D reconstruction of organoids were made with Blender, a free and open-source 3D computer graphics software. Each cell was represented by a 3D sphere and could be colored based on the (inferred) cell types (**figure 4.1B, F**).

Organoid fixation and permeabilization. Organoid samples were fixed with 4 % formaldehyde (Sigma-Aldrich) at room temperature. In order to get rid of the gel but keep the organoids attached to the plate, we optimized the fixation protocol. After adding formaldehyde, we waited for ~ 10 minutes and then gentling washed the sample with PBS to remove the gel, which otherwise would hinder the penetration of antibodies and reduce imaging quality. 10 minutes was the optimized waiting time to ensure gel removal and more than 50 % of the imaged organoids attached to the cover glass (**supplementary figure 4.1C**). After gel removal, organoid samples were incubated in formaldehyde again for 20 minutes to complete the fixation procedures. Following fixation, permeabilization was performed by incubating the samples in 0.2 % Triton-X-100 (Sigma-Aldrich) for one hour at room temperature. All the washing procedures were performed gently to avoid removing organoids from the cover glass.

Staining with antibodies and dyes. Following fixation and permeabilization, organoids were blocked with 5 % skim milk in TBS at room temperature for one hour. Subsequently, organoids were incubated in blocking buffer containing primary antibody (see section antibodies) for two days at 4 °C, and then incubated with secondary antibody (see section antibodies) at room temperature for one hour. These procedures would be repeated for each antibody. Regarding the dyes, organoids were incubated with Wheat germ agglutinin (WGA) conjugated to CF©488A (5 ug/ml Biotium) at room temperature for two hours and with RedDot™1 Far-Red Nuclear stain (1:200 Biotium) or SYTOX™ Orange Nucleic Acid stain (1:5000 Thermo Fisher Scientific #S 11368) at room temperature for 20 minutes.

Antibody stripping. After imaging the results from each round of antibody staining, the primary antibodies were removed by incubation with elution buffer at room temperature for 15 minutes while shaking[46]. This was repeated six times with the elution buffer replaced between consecutive cycles. The elution buffer was prepared by adding 0.5 M Glycine (Sigma-Aldrich), 5 M Urea (Sigma-Aldrich), 5 M Guanidinium chloride (Sigma-Aldrich), 70 mM TCEP-HCL (Sigma-Aldrich) to H_2O, with pH adjusted to 2.5.

Antibodies and dyes. The primary antibodies used in this study are: Rabbit anti-Lysozyme (1:800, Dako #A0099), Rabbit anti-Olfm4 (1 : 500, Cell signaling technology #39141), Recombinant Rabbit anti-Aldolase B + Aldolase C (1:300, Abcam #ab75751), Mouse anti-Human Cytokeratin 20 (1 : 500, Dako #M701929-2), Mouse anti-Chr-A (1 : 50, Santa Cruz Biotechnology #sc-393941), Rat anti-E-cadherin (1:400, Santa Cruz Biotechnology #sc-59778), CD24 Monoclonal Antibody (1:200, ThermoFisher #17-0242-80), Mouse anti-Ki67

(1:200, BD Biosciences 550609). The secondary antibodies used in this study are: Goat anti-Rabbit IgG H&L (Alexa Fluor©405) pre-adsorbed (1:1000, Abcam #ab175654), Goat anti-Rat IgG H&L (Alexa Fluor©555) pre-adsorbed (1:1000, Abcam #ab150166), Donkey anti-Mouse IgG H&L (Alexa Fluor©647) (1:500, Thermo Fisher #A31571), Donkey anti-Rabbit IgG H&L (Alexa Fluor©405) pre-adsorbed (1:1000, Abcam #ab175649). The dyes used in this study: WGA conjugated to CF©488A (5 ug/ml, Biotium), RedDot™1 Far-Red Nuclear stain (1:200, Biotium), SYTOX™ Orange Nucleic Acid Stain (1:5000, Thermo Fisher Scientific #S11368). The order of staining, optimized to ensure good staining quality for all cell types, was based on the staining quality and stripping difficulty of each antibody (**supplementary figure 4.1G**).

Mapping endpoint cell types to lineages. During time-lapse imaging, the Leica software allowed recording of the imaged locations, which could thus be found back after imaging. To achieve this, the mounting stage of the microscope and the orientation of the cover glass should be consistent with the settings during time-lapse imaging. With our optimized protocols for sample preparation and fixation (see Section Organoid fixation, permeabilization and staining), we could keep more than 50 % of the imaged organoids with limited deformations in the plate after fixation. After staining and relocating the organoids that were imaged in time during growth, mapping all cells (including their type information) to the cells that were tracked could be still challenging, due to the constant movement of cells during growth and global rotation and deformation of organoids caused by the fixation and repeated staining and washing. To mitigate these issues, we fixed the organoids within 5 minutes after the time-lapse imaging and performed every washing step gently, in order to preserve the spatial context of single cells. The linking of cells before and after fixation could be achieved mostly based on the spatial context of each cell. Linking was done in two steps. In a first automated step, we used a min-cost flow solver algorithm[99], which integrally optimizes the linking for all the tracked cells, and was also employed for the similar task of tracking cells between frames during organoid growth. In a second step, we manually corrected the automated linking results by visual inspection of the movie and staining images. The fluorescence intensity of H2B-mCherry showed heterogeneity between cells during time-lapse imaging, which could be preserved during fixation. Therefore, the brightness of the nuclear marker could also assist cell linking before and after fixation.

Endpoint sister type analysis. For cells present at the endpoint, identified as specific type and with a sister, we checked the possible cell types of the sister pairs and counted the occurrence of each combination. For each cell type, we counted the number of sister pairs where at least one of them was of that type. If both sisters were of that type, the pair would be counted twice. Cells of different types had different abundance and majority of the sister pairs contained stem cells and/or TA cells. We then normalized the 2D histogram, via dividing the occurrence of each combination by the sum of each column, as shown in the bar

plot in **figure 4.2A**. Therefore, the frequency within each column in **figure 4.2A** would sum up to be 1. The analysis was based on nine different organoids.

Cell age distribution analysis. For cells present at the endpoint, the duration between the birth time and the endpoint of imaging was measured as the cell age. Some cells were present from the beginning till the end. Their ages were then measured by the total length of the imaging experiment duration (~ 60 hours). The age distribution of each cell type was studied and plotted as a box plot in **figure 4.2B**, followed by statistical significance tests. Box plot elements represent the following: center line: median; box, quartiles; whiskers, range; fliers, outliers. The analysis was based on seven different organoids from experiments lasting ~ 60 hours.

Analysis within (sub-)trees containing two cell types. All the (sub-) lineage trees with two different cell types were taken into account, unless more than 50 % of the cells within the lineage could not be tracked or died. The absolute count of each possible combination of the two cell types was shown in the 2D histogram (**supplementary figure 4.10**).

Cell type backpropagation. The assumption underlying the backpropagation of cell types is that changes in cell types are rare. This assumption could be supported by the found type symmetry between sisters since frequent type changes likely lead to different types in sisters. Starting at the lineage endpoints, we propagate the measured endpoint types back in time following this process:

Backpropagation along consecutive timepoints. From one timepoint to a previous timepoint, the type is initially assigned as unchanged if no tree branch points (divisions) is traversed (marked '1' in **figure 4.2C**).

Backpropagation of symmetric fate. For cell types identified by endpoint staining, we observed that sisters almost always assumed the same fate (**figure 4.2A**), suggesting that this fate was already set in the mother cell. Generalizing this observation, we assumed that if both sisters have the same (inferred) cell type, the inferred cell type of the mother cell is the same (marked '2' in **figure 4.2C**). Regarding cells with a dead sister, the mother is inferred the same type as the living daughter.

Backpropagation of asymmetric fate. The above backpropagation rule does not apply if two daughters have different (inferred) cell types. Therefore, we introduced two addition backpropagation rules. First, if at least one daughter's (inferred) cell type was stem cell, then the inferred cell type of the mother was also stem cell (marked '3' in **figure 4.2C**). Second, if the (inferred) cell type of one daughter was TA and the other daughter was not stem cell, then the inferred cell type of the mother was TA (**figure 4.2D**). These two rules were based on the capability of stem cells to generate all cell types and the transient property of TA cells between stem cells and differentiated cells.

Forward propagation of cell type changes. If a mother and a daughter cell had different (inferred) cell types, we interpreted this as a change in cell type that occurred during the lifetime of the daughter cell (marked with triangle in **figure 4.2C, D**).

These simple rules were sufficient to propagate backwards the lineage trees that we have encountered, unless the tree appeared very 'broken' where majority of the cells could not be tracked or died. All the lineage trees after backpropagation were shown in **supplementary figure 4.7.**

Imaging of a Lgr5 reporter. To test our backpropagation method, we performed time-lapse imaging, endpoint staining and live-cell tracking in an organoid line with both Lgr5-GFP, a well-known stem cell marker, and a H2B-mCherry reporter. To limit phototoxicity caused by GFP imaging, the time-lapse imaging lasted for around 24 hours, followed by endpoint staining. Quantification of the membrane bound Lgr5-GFP fluorescence signals was conducted by determining the average fluorescence intensity within a 2D circle centered around the membrane area with a diameter of 6 μm (**supplementary figure 4.8A**). The Olfm4 staining was measured with the same method.

Imaging of Paneth cells and Enteroendocrine cells. To image Paneth cells, CD24 antibody was added to the medium of H2B-mcherry organoids one hour before imaging. Time-lapse imaging was performed in both H2B-mCherry and CD24 (647nm) channels for more than 24 hours. To image Enteroendocrine cells, a mouse organoid line carrying a Neurogenin 3-tdTomato reporter was used[147].

Measuring locations of cells along the crypt-villus axis. At each timepoint, the crypt-villus axis was manually annotated in the xy plane at the z position corresponding to the center of the crypt, since tracked crypts grew perpendicularly to the objective. Three to six points were marked along the axis, through which a spline curve was interpolated as the axis. For each tracked cell i we determined its position along the spline by finding the value of r_i that minimized the distance d between the cell position and the axis (**supplementary figure 4.12A**). The bottom-most cell of the crypt, i.e. that with the lowest value of r_i, was defined as position zero. Based on the shape and curvature of the epithelium, the location of crypt neck (where there was a sharp transition from crypt to villus) was estimated and annotated manually, as an indication of the length of the crypt. Since different crypts were of various length, we did a normalization of the locations based on the crypt neck location. For each cell's measured distance in μm within a certain frame, we divided it by the distance from the crypt neck within the same frame to the crypt bottom. Therefore, the length from crypt neck to bottom would remain one for each timepoint and each crypt. With this measurement, both the locations of different (inferred) cell types and the type transitions were mapped along the crypt-villus axis (**figure 4.3A-E**).

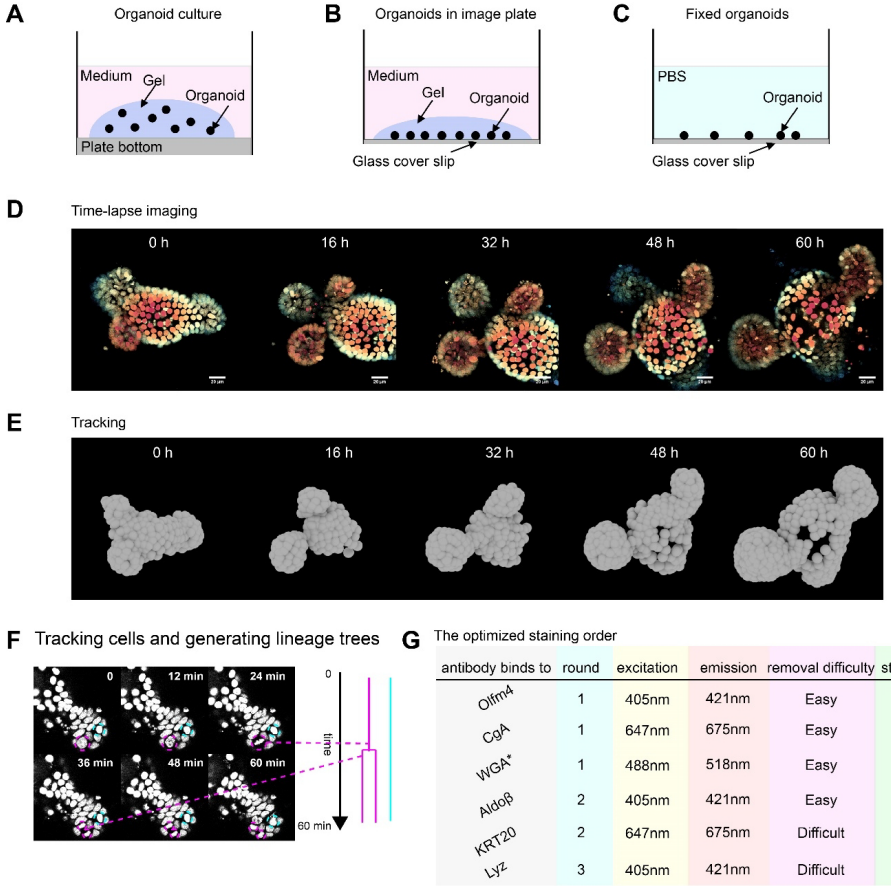
Measurement of migration speed along the axis. To estimate how fast a cell migrated along the axis, we searched for the locations of cells along the crypt-villus axis when they firstly

showed up during tracking and the locations of cells when they were last present. The migration speed could be estimated by dividing the distance that the cell had migrated by the duration during which the cell was present (**figure 4.3F**).

Search for neighbors for each cell. Cells could have varying numbers of neighbors because of the disorder in the epithelium. Distances between nuclei could vary between cell types and location (spread apart in the villus-like region and closely packed in the crypts). To obtain robust neighbor pairs, we functionally defined neighbors as pairs of nuclei without another nucleus in between (**supplementary figure 4.12D**). This condition was tested by a 'neighbor score', the ratio of the sum of the distances of the two cells of interest (A&B) to a third cell (S) and the distance between the two cells (A&B), namely $\frac{d_{AS}+d_{BS}}{d_{AB}}$. If the third cell S positioned perfectly in between the pair of interest A & B, the neighbor score would appear as the minimal value of 1, and A and B would not be identified as neighbors. If A and B were not separated by S, the three nuclei would form a triangle with high neighbor score between A and B. For each cell, we calculated the neighbor score for the twenty closest neighbors (in Euclidean distance) at every timepoint. If the neighbor score were higher than $\sqrt{2}$, we would consider them neighbors. This cut-off corresponded to diagonal neighbors in the case of a perfect square lattice. Using this cut-off, we found most cells with five or six neighbors, exactly as expected for the basal side of a curved epithelium[186].

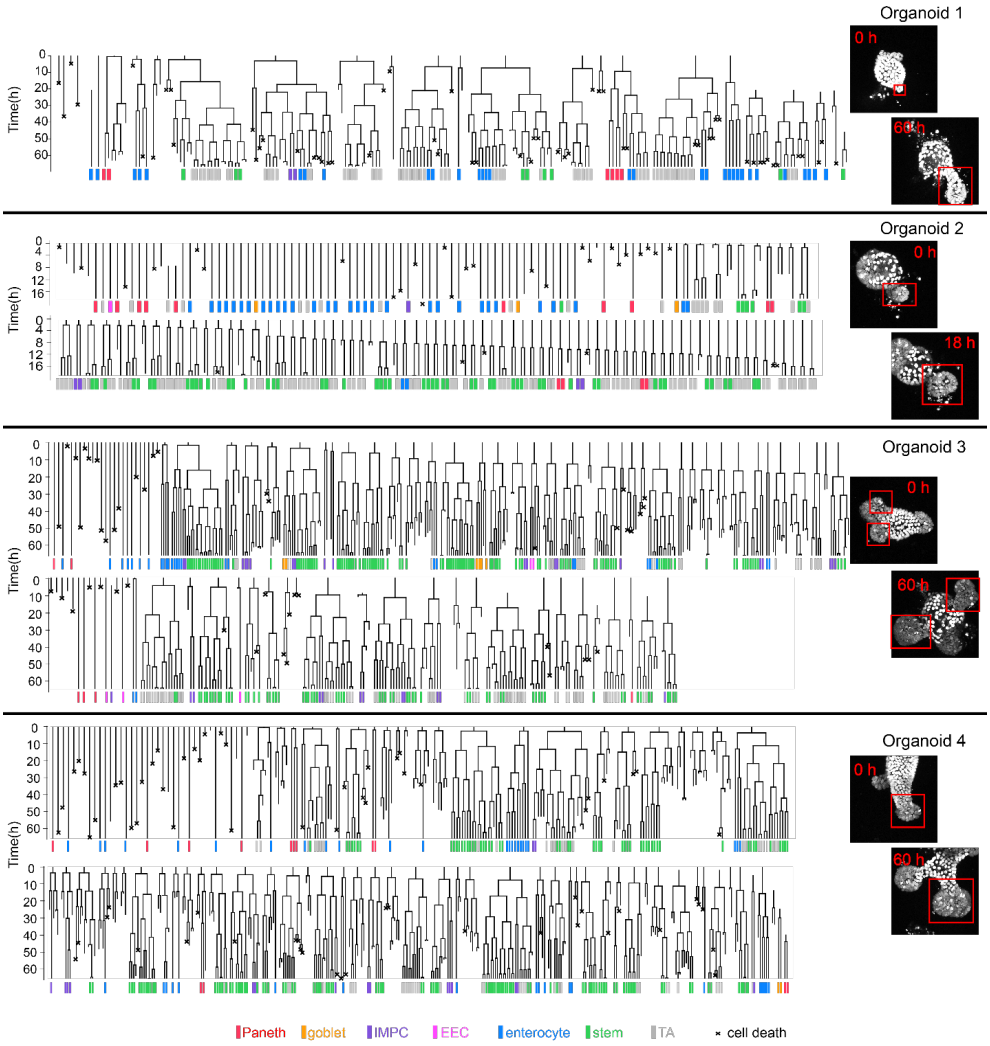
Measurement of separation rate. Separation rates were determined by following pairs of neighbors over time. For a newly born cell, its neighbors were searched and selected with the method introduced above. The selection was conducted one hour after division, so that the nuclei would have returned to the basal side of the epithelium. If the selected neighbors divided, we would continue tracking one of the daughters (selected randomly) so that the following of the neighbor pairs would not be cut short by division. The separation rates were measured after following the neighbor pairs for 2 hours and 10 hours, by calculating the fraction the pairs staying as direct neighbors within the total of pairs that were followed. Regarding the rearrangement rates of sisters, we followed the sister pairs that shared the same (inferred) cell type. Separation rates of sister pairs were also measured after 2 hours and 10 hours, by calculating the fraction of sisters staying as direct neighbors within the total number of sister pairs being followed (**figure 4.3H, I**).

4.5 Supplementary figures

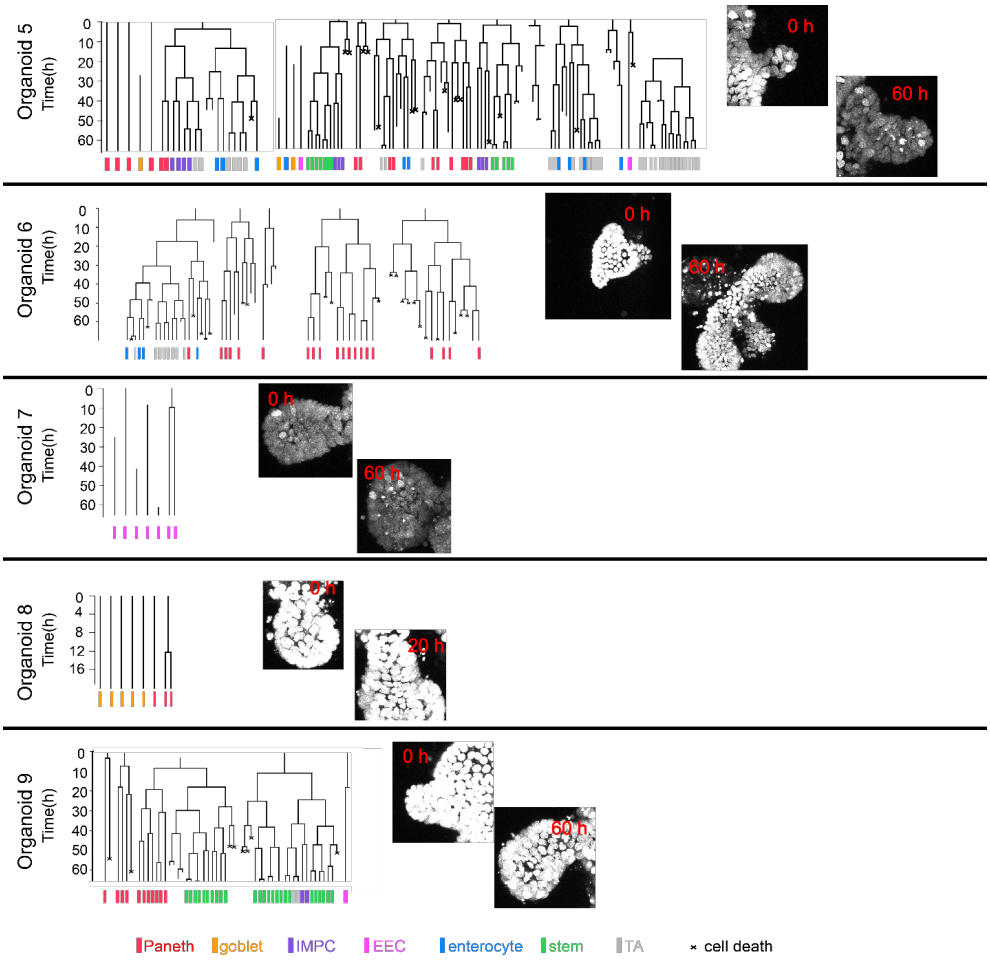


antibody binds to	round	excitation	emission	removal difficulty	staining signal
Olfm4	1	405nm	421nm	Easy	Dim
CgA	1	647nm	675nm	Easy	Good
WGA*	1	488nm	518nm	Easy	Good
Aldoß	2	405nm	421nm	Easy	Good
KRT20	2	647nm	675nm	Difficult	Good
Lyz	3	405nm	421nm	Difficult	Good

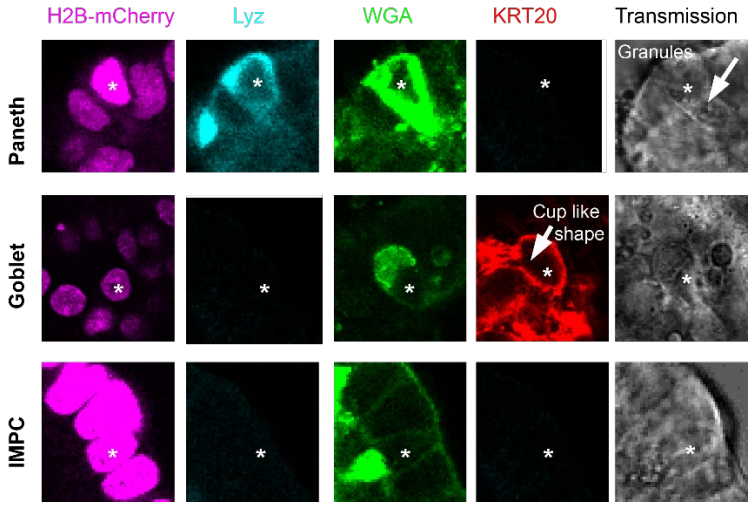
Supplementary figure 4.1: TypeTracker applied to mouse intestinal organoids. **A,** For conventional organoid culture, organoids were scattered in 'domes' of BME gel and located at various heights. **B,** For time-lapse imaging, organoids were seeded in a thin layer of BME gel in chambered cover glass slides. Immediately after seeding, samples were put in fridges for 10 minutes so that organoids all sank towards the cover glass. **C,** With our optimized protocol, more than 50 % of the imaged organoids would remain at their imaged locations after fixation. **D,** Time-lapse imaging of an organoid carrying the H2B-mCherry reporter with 3D confocal for 60 hours. Scale bar, 20 µm. Color encodes different z-planes. **E,** Live-cell tracking of the organoid. In these 3D reconstructions, each cell was represented by a sphere centered at the estimated nuclear center. Cells were not tracked if they were located far away from the objective and would move away from the region of interest. **F,** Tracking cell divisions and generating lineage trees. Cells were tracked between consecutive frames based on their locations, such as the cell highlighted by the cyan circle. Cell division events could be identified by the morphology changes of the nuclei (highlighted with the magenta circle) and recorded with the branching of lineage trees shown in magenta. **G,** The order of antibodies and dyes to use in different rounds was optimized based on each antibody's staining quality and stripping difficulty.



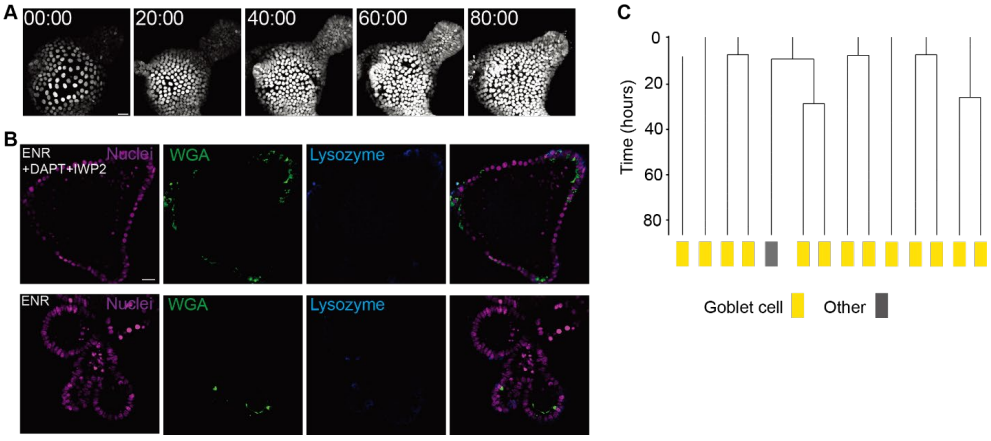
Supplementary figure 4.2: Gallery of lineage trees, generated from live-cell tracking, with cell types mapped at the endpoint. These trees are from four different organoids, of which the max projections of H2B-mCherry signals at the beginning and the end of time-lapse imaging are shown. The regions of tracked cells are marked by red boxes.



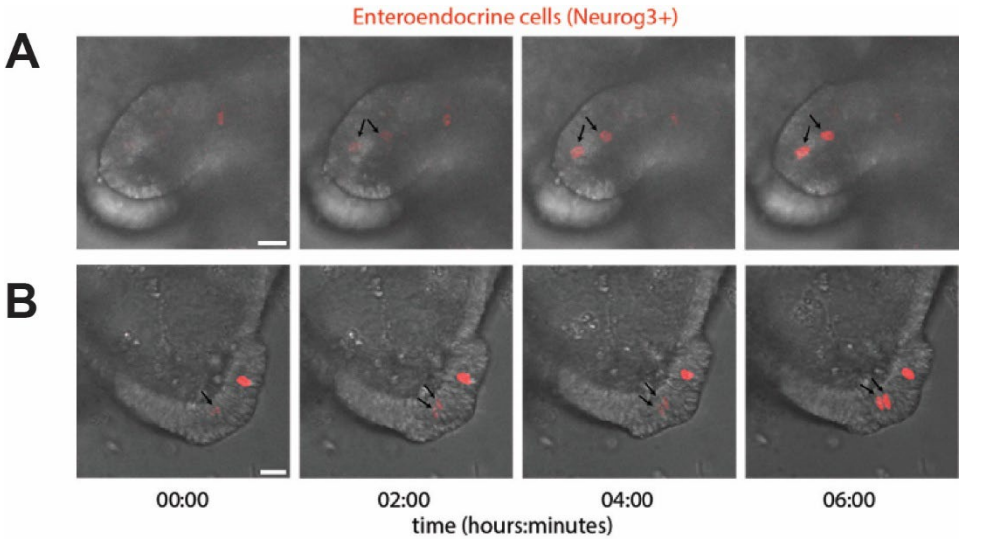
Supplementary figure 4.3: Gallery of lineage trees from organoids where certain cell types were tracked. The max projections of H2B-mCherry signals at the beginning and the end of time-lapse imaging are shown.



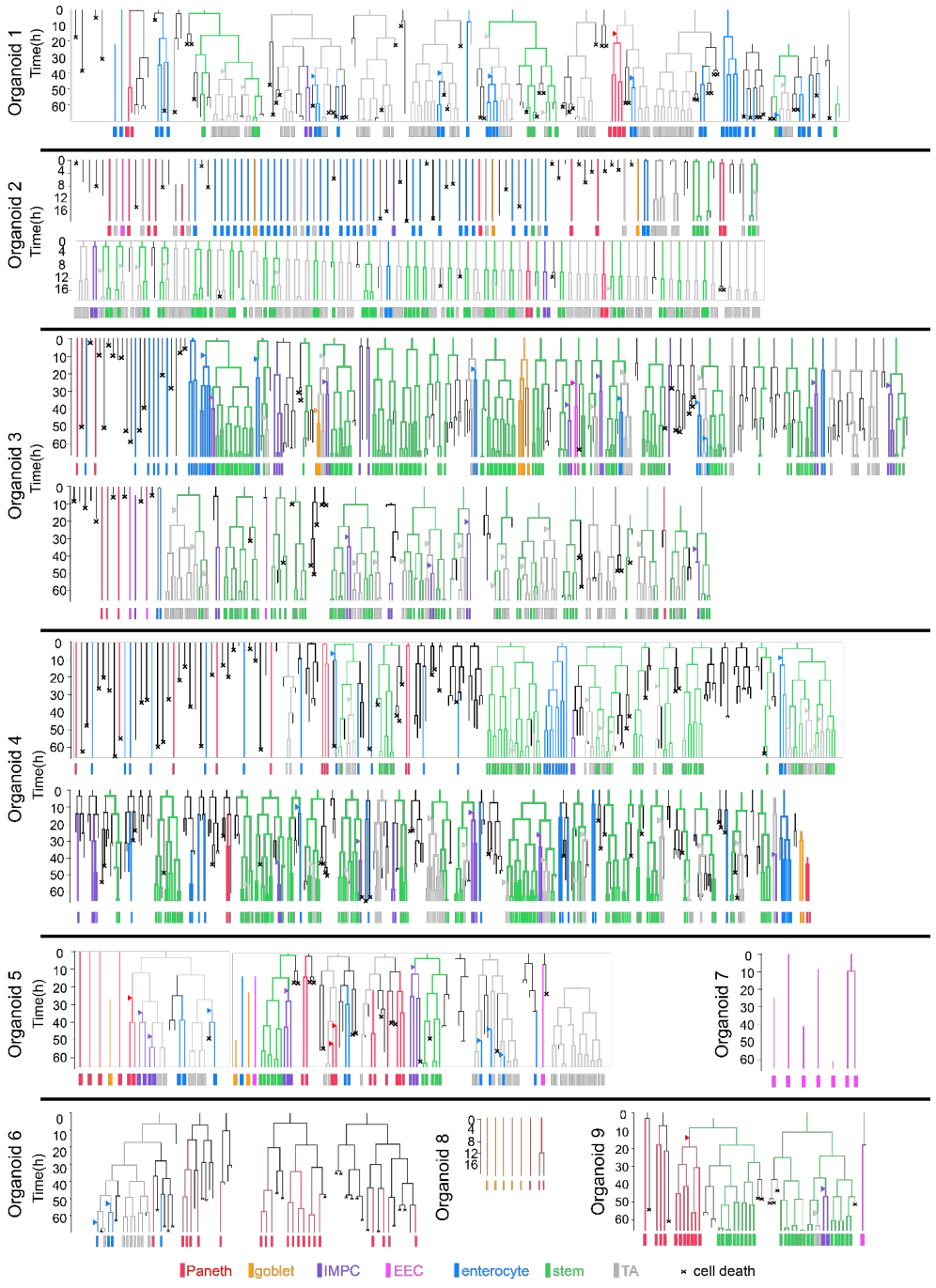
Supplementary figure 4.4: The identification of mature Paneth cells and goblet cells. Paneth cells (indicated with * in row 1) often showed extremely bright H2B-mCherry fluorescence signals compared with the neighbor cells, bright Lysozyme (Lyz) fluorescence signals at the basal side of the cell, bright Wheat Germ Agglutinin (WGA) staining and granules in the transmission channel. Goblet cells (indicated with * in row 2) often stained positive of Cytokeratin 20 (KRT20) and WGA, with a cup-like shape. A group of cells that stained positive of WGA but negative of KRT20 or Lyz were called Immature Mucus producing cells (IMPCs, indicated with * in row 3). These cells could be early Paneth cells or goblet cells considering the mucus secretion functions that they had.



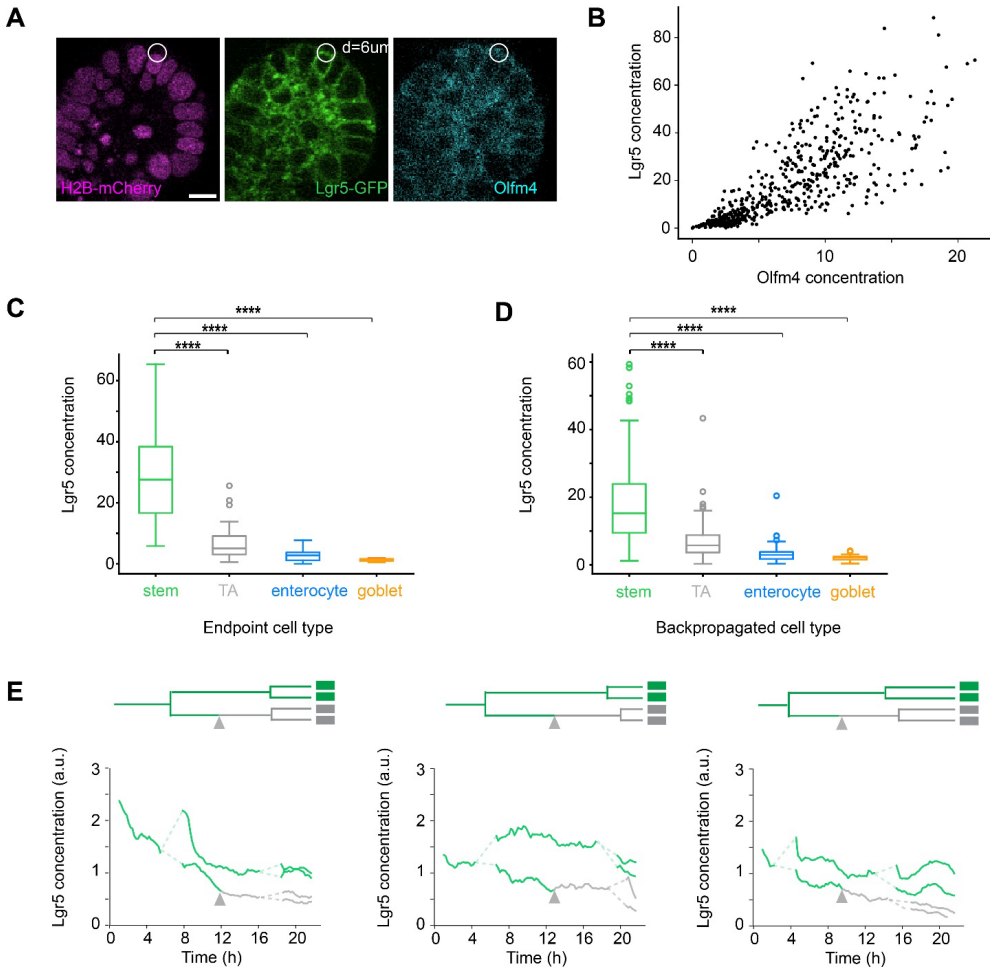
Supplementary figure 4.5: Sister symmetry for goblet cells in enriching conditions. **A,** Time-lapse imaging of organoids in culture media supplemented with IWP2 and DAPT for 80 hours. **B,** Staining results. **C,** TypeTracker lineages of cells identified as goblet cells at the endpoint by staining. The data are consistent with sister symmetry for secretory types, and early commitment in the mother cell.



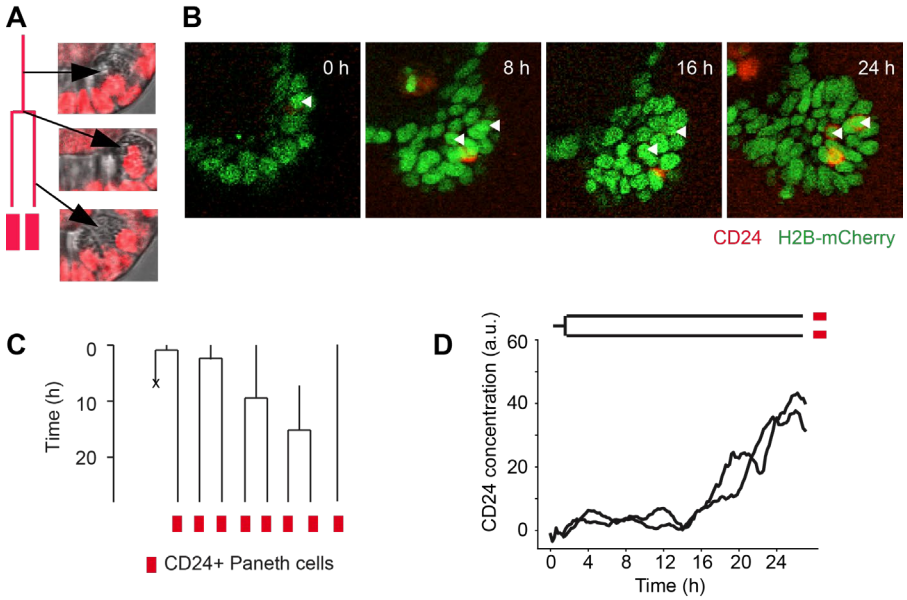
Supplementary figure 4.6: Sister symmetry for enteroendocrine cells visualized with a Neurogenin 3 reporter. Time-lapse imaging of organoids carrying a Neurogenin 3 reporter. A and B show two examples where cells (marked by arrows) located close to each other started to express Neurogenin 3 signals at a similar time. The data are consistent with sister symmetry for secretory types, and early commitment in the mother cell.



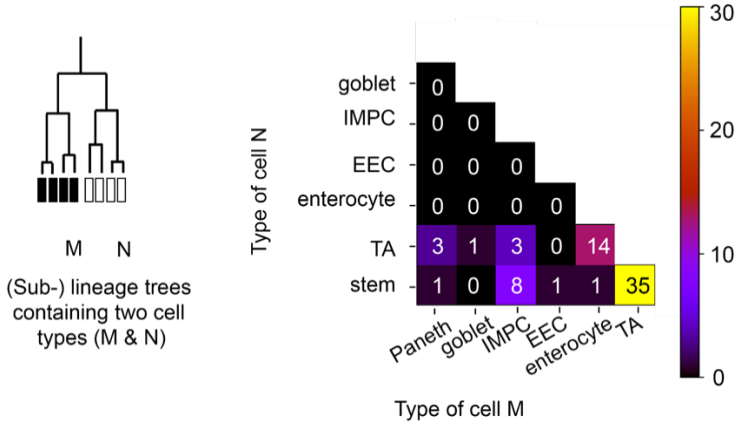
Supplementary figure 4.7: Gallery of lineage trees after backpropagation with inferred cell type transitions. Inferred cell types are shown with different colors. Type transitions are indicated by triangles, colored based on which type the cell was inferred to transition towards. These trees are from nine different organoids.



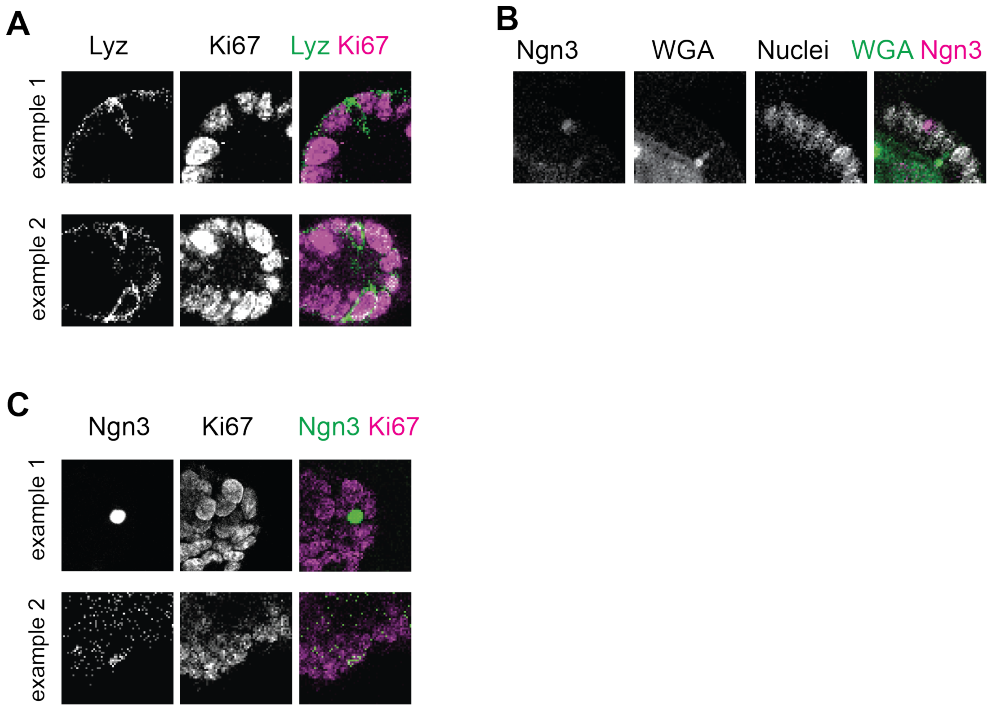
Supplementary figure 4.8: Imaging and tracking with an Lgr5 reporter confirmed the backpropagation method. A, Images of the H2B-mCherry (nuclei), Lgr5-GFP (fluorescent live reporter of stem cells), and Olfm4 (antibody used to identify stem cells). In order to estimate the fluorescence concentration at each cell's membrane, the fluorescence within a 2D circle (d = 6 µm) at the membrane region was measured. **B,** The Lgr5 and Olfm4 fluorescence concentrations, measured by averaging the fluorescence intensity within 2D circles as shown in A, were proportional in single cells. **C,** At the endpoint, the Lgr5 fluorescence concentration of single cells plotted against cell types identified with antibody staining. **D,** The Lgr5 fluorescence plotted against cell type inferred by the TypeTracker method, showing similar distributions to the plots at the endpoint in C. **E,** Lgr5 fluorescence in lineages going through transitions from stem cells to TA cells. Consistently, Lgr5 is lower in lineages that make this transition, compared to the cousin lineages that retain the stem fate. Moreover, the Lgr5 decrease is indeed established in the mother cell that was identified as the committing cell by the TypeTracker method (triangle), thus validating the latter.



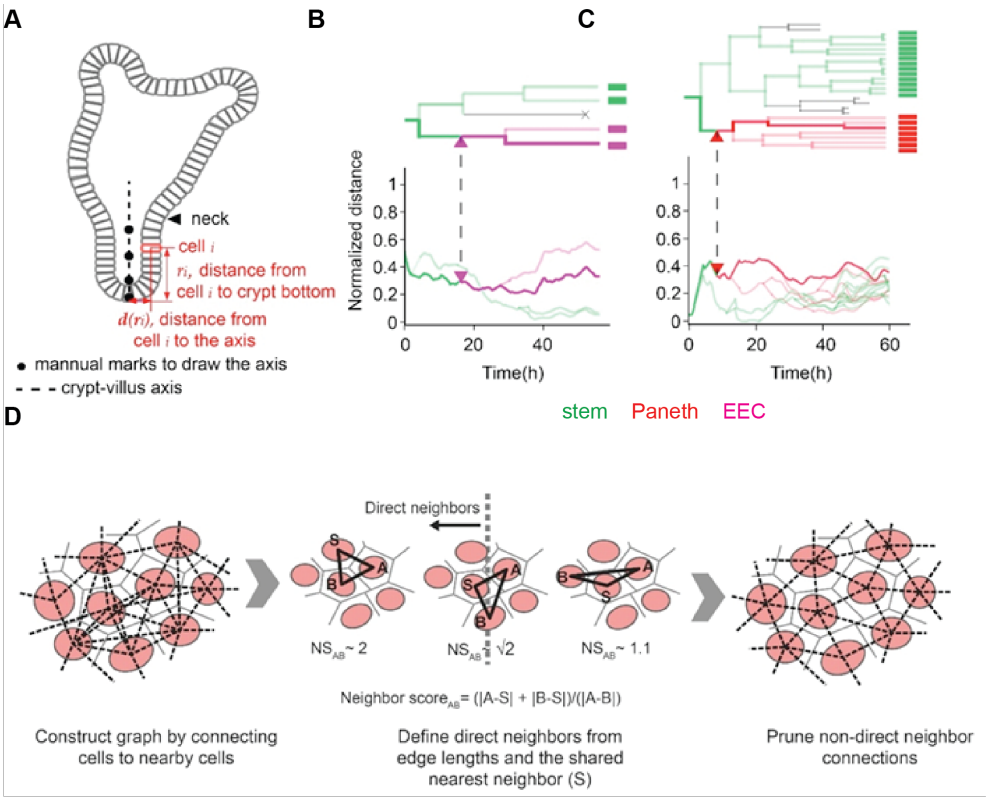
Supplementary figure 4.9: Sister symmetry and real-time marker quantification for Paneth cells. A, Paneth cell divisions. Granules are seen before and after division, evidencing Paneth cell division. **B,** Time-lapse imaging of H2B-mCherry (nuclei) and CD24 (antibody marking Paneth cells). The white arrow indicates a CD24- cell dividing into two cells which both became CD24+ later during imaging. **C,** Lineage trees of CD24+ cells. **D,** The fluorescence signals of CD24 are plotted with time in the example lineage in B.



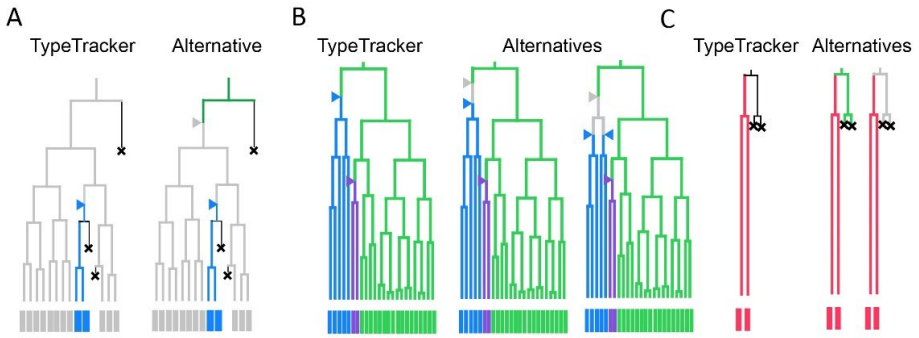
Supplementary figure 4.10: (Sub-) lineage trees containing two cell types often had at least one of the types as either stem cells or TA cells. All the (sub-) lineage trees with two different cell types were taken into account unless more than 50 % of the cells within the lineage could not be tracked or died. The occurrence of each possible combination of the two cell types was counted and shown in the 2D histogram. A combination of two different differentiated cell types, such as enterocytes and goblet cells, was never found. Differentiated cells were often found together with either stem cells or TA cells in the same lineage.



Supplementary figure 4.11: Co-staining cell type markers with Ki67. A, Most Lysozyme (Lyz) positive cells do not co-stain with the proliferation marker Ki67 (example 1, 69%, N = 105 cells), but some do (example 2, 31%, N = 105 cells), supporting the idea that Paneth cells can divide after commitment. **B,** Neurogenin 3 (Ngn3) positive cells do not co-stain with Wheat Germ Agglutinin (WGA, N = 16 cells). **C,** Most Ngn3 positive cells do not co-stain with Ki67 (example 1, 81%, N = 16 cells), but some do (example 2, 19%, N = 16 cells), supporting the idea that Enteroendocrine cells can divide after commitment.



Supplementary figure 4.12: Following the spatial organization of cells during differentiation. A, The crypt-villus axis could be generated by interpolating through the manually annotated points. For each tracked cell i, we determined its position along the axis by finding the value of r_i that minimized the distance d(r_i) between the cell position and the axis. **B & C,** The moving trajectories of cells within different lineages were colored by inferred cell types. Transitions to EECs and Paneth cells took place deep in the crypt, around 0.4, surrounded by stem cells which were often found from 0 to 0.6 along the axis. **D,** Neighbors were defined as pairs of nuclei without another nucleus in between. Each cell's neighbor score for the twenty closest cells (in Euclidean distance) was calculated at every time point. If the neighbor score were higher than √2, cells would be identified as neighbors.



Supplementary figure 4.13: TypeTracker alternative type assignments and potential errors. A, Example lineage tree where the stem and TA fates are ambiguous early in the tree. In this tree, with TA but without stem endpoint fates, the founding cell type may be TA (grey), as TypeTracker would predict, or stem (green) whose descendant lineages are all committed to TA or other types (Alternative). High cell death rates (crosses) exacerbate this issue, as aborted lineages cannot propagate type information back in time. Overall, this problem is mitigated by the fact that proliferative stem and TA types generate multiple branches, which helps their type maintenance in at least one lineage until the end of the tree, and thus proper inference. **B,** Example lineages where a potential TA stage could be missed. **C,** Interrupted lineages. Lineages that are interrupted because of cell death or tracking errors do not provide type information. For instance, in the example shown one then cannot infer the fate of the mother cell that founded the two Paneth cell sisters. This error does not involve an identification ambiguity or error per se, as the lack of information is clear. However, it does limit the backpropagation of genealogically related lineages that are type-identified, as the type of the mother cell then remains undetermined. Such lineages are identified by their black colour in the trees.

5 Sequential expression of tuft cell subtypes along the crypt-villus axis

Intestinal tuft cells are epithelial sentinels that trigger host defense upon detection of parasite-derived compounds. While they represent potent targets for immunomodulatory therapies in inflammation-driven intestinal diseases, their functioning and differentiation are poorly understood. Here, we reveal common intermediary transcriptomes among the previously described tuft-1 and tuft-2 subtypes in mouse and human. Tuft cell subtype-specific reporter knock-ins in organoids show that the two subtypes reflect successive post-mitotic maturation stages within the tuft cell lineage. In vitro stimulation with interleukin-4 and 13 is sufficient to fuel the generation of new Nrep+ tuft-1 cells, arising from tuft precursor (tuft-p) divisions. Subsequently, changes in crypt-villus signaling gradients, such as BMP signaling, are required to advance maturation towards Chat+ tuft-2 phenotypes. Functionally, we find chemosensory capacity to increase during maturation. Our tuft subtype-specific reporters and optimized differentiation strategy in organoids provide a platform to study immune-related tuft cell subtypes and their unique chemosensory properties.[1]

[1] The content of this chapter is published as: Buissant des Amorie, J.R., Betjes, M.A., Bernink, J.H. et al. Intestinal tuft cell subtypes represent successive stages of maturation driven by crypt-villus signaling gradients. *Nat Commun* 16, 6765 (2025).

5.1 Introduction

Tuft cells are solitary chemosensory epithelial cells that respond to specific environmental stimuli by secretion of effector molecules that initiate appropriate changes in the immune state and physiology of their surroundings. Although they are rare, they attract significant interest now that increasing amounts of evidence underscore their crucial role at the interface between the intestinal epithelial barrier and the immune system, making them interesting targets for therapeutic immunomodulation[187-189]. In the intestine, tuft cells detect parasitic helminths and *Tritrichomonas* protists and secrete the cytokine interleukin-25 (IL-25) to initiate the type 2 immune response required for parasitic clearance, characterized by increased mucus production (*weep*) and muscle contractility (*sweep*)[190-192]. More specifically, IL-25 activates innate lymphoid cells (ILC2s) in the lamina propria and triggers the release of type 2 cytokines, including interleukin-13 (IL-13). IL13, together with IL4, acts on epithelial progenitors and biases their differentiation towards goblet and tuft cells, resulting in a tuft-ILC2 feedforward circuit that leads to a strong increase in tuft cell numbers, up to as much as 8% in the case of a helminth infection[190].

In addition to remodeling of the intestinal epithelium through the tuft-ILC2 circuit, tuft cells also signal directly to nearby epithelial cells. For instance, tuft cells are the only epithelial cells that express choline acetyltransferase (ChAT), the enzyme catalyzing acetylcholine biosynthesis. In turn, tuft cell-derived acetylcholine induces fluid secretion from enterocytes by stimulating apical Cl- ion release, further contributing to the '*weep'* into the intestinal lumen[193]. Similarly, prostaglandin D2 secreted by activated tuft cells can increase mucus secretion from goblet cells[194].

Like most epithelial cell types of the small intestine, tuft cells arise from *Lgr5*+ stem cells at the crypt base and migrate up towards the tip of the villus where they are eventually shed into the intestinal lumen[195,196]. Following the discovery of tuft cells as important sentinels and signal relays, multiple studies have shed light on the mechanisms of tuft lineage specification. For instance, the transcription factor POU class 2 homeobox 3 (POU2F3) was found to be a master regulator of tuft cell fate and other transcription factors, such as Gfi1b and Sox4, were found to be important for tuft lineage commitment[192,197,198]. Interestingly, while loss of Notch signaling followed by expression of atonal bHLH transcription factor 1 (*Atoh1*) is essential for differentiation towards most secretory cell types (goblet, enteroendocrine and Paneth), the role of Notch signaling and *Atoh1* in tuft cell differentiation seems to be more complex. Multiple studies describe *Atoh1*-independent tuft cell populations in the small intestine, while there is evidence for *Atoh1*-dependent populations in the colon[195-199].

Despite these insights into early tuft cell fate specification, much less is known about the signals that further advance tuft cell differentiation and maturation, nor how functionality relates to transcriptional heterogeneity that was observed within the intestinal tuft cell pool[17]. Indeed, two different types of tuft cell have been identified by single-cell RNA sequencing: tuft-1 cells with a neuron-related transcriptional program and tuft-2 cells with an immune-

related transcriptional program. Computational analysis of spatial transcriptional programs along the crypt-villus axis, suggested zonation of these two tuft types with tuft-2 markers expressed higher up towards the tip of the villus and tuft-1 markers towards the bottom of the villus[15]. While a tuft-1 cell specific function in the small intestine has yet to be found, sensing of the bacterial metabolite N-undecanoylglycine through vomeronasal receptor Vmn2r26, has been mainly attributed to tuft-2 cells[194]. Although multiple metabolite GPCRs, such as *Ffar3* and *Tas1rs*, as well as orphan GPCRs, are specifically expressed by intestinal tuft cells[17], tuft cell activation by ligand stimulation has only been established for the succinate receptor SUCNR1 and the aforementioned vomeronasal receptor Vmn2r26[194,200].

Here, we analyze single-cell intestinal tuft cell transcriptomes from mouse and human to reveal conserved proliferative tuft cell precursors (tuft-p), as well as abundant intermediaries of the post-mitotic tuft-1 and tuft-2 cell states. Moreover, new organoid models with tuft cell subtype-specific reporter knock-ins show that treatment with IL-4 and IL-13 leads to increased tuft cell numbers, but solely generates tuft-1 states. Adaptation of the crypt-like organoid medium to better mimic the villus environment, where tuft-2 cells reside *in vivo*, facilitates transitioning of post-mitotic immature tuft-1 states to mature tuft-2 states. Practically, our optimized organoid differentiation strategy provides tuft cells that closely resemble their functional counterparts *in vivo* and are amenable for in-depth cell biological studies, such as real-time imaging of tuft cell morphological and chemosensory properties.

5.2 Results

5.2.1 Gradual Chat expression marks a continuum of zonated tuft cell states on the crypt-villus axis

To characterize intestinal tuft cell heterogeneity with high-resolution, we performed single-cell RNA sequencing on a FACS-enriched tuft cell population isolated from *Chat^{BAC}-eGFP* mice where tuft cells are labeled with enhanced green fluorescent protein (eGFP), based on tuft cell-specific epithelial expression of choline acetyltransferase (*Chat)*[201](**Figure 5.1a; Supplementary Figure 5.1ab**). Unsupervised clustering resulted in three clusters (**Figure 5.1a**), with the two largest identified clusters showing high similarity to the previously described tuft-1 and tuft-2 populations[17] (**Figure 5.1b**). The third and smallest cluster robustly expressed general tuft cell markers such as *Dclk1*, *Trpm5* and *Avil,* but did not show strong expression of genes related to tuft cell chemosensory or immunological function such as *Sucnr1, Chat*, and *Il25* (**Figure 5.1c**). Differential expression analysis indicated that this cluster separated from the other two clusters based on gene expression that is also detected in other cellular lineages, such as *Aldob* and *Dmbt1* (enterocytes), *Muc3* (Goblet cells) or *Reg3b* and *Reg3*g (Paneth cells), leading to a relatively low tuft-specific signature score (**Supplementary figure 5.1cd**). Moreover, components required for MHCII-dependent antigen presentation, such as *Cd74*, *H2-Ab1* and *H2-Aa*, earlier found to be expressed by

intestinal stem cells[202], were expressed in this cluster (**Supplementary figure 5.1d**). Last, we observed high expression of proliferation markers *Mcm2*, *Mki67*, *Top2a* and *Pcna* (**Figure 5.1c**) and a large proportion of this cluster was predicted to be in S/G2/M phase of the cell cycle (**Supplementary figure 5.1e**). Collectively, these data suggest that the third cluster represents immature early-state tuft precursors (tuft-p). We validated the existence of the three transcriptomic tuft cell states by integration with other scRNA-seq datasets of the mouse small intestinal epithelium[17,173,203] (**Figure 5.1d**). Affirmatively, we found tuft-p, tuft-1 and tuft-2 clusters among the more than 2000 tuft cell profiles in the integrated dataset, supported by differential expression of their respective transcriptional signatures (**Figure 5.1e; Supplementary figure 5.1f**).

Next, we evaluated *Chat* expression within the tuft cell pool and noticed substantial heterogeneity, while general tuft cell marker *Dclk1* was uniformly expressed (**Figure 5.1f**). Specifically, *Chat* was most prominently expressed in tuft-2 cells, which is corroborated by differential expression levels in our scRNA-seq dataset of sorted *Chat*^{BAC}-eGFP+ cells (**Figure 5.1g**). We validated heterogeneous *Chat* expression within the tuft cell pool by immunofluorescence against DCLK1 and GFP on intestinal tissue sections of the *Chat*^{BAC}-eGFP mice. This confirmed *Chat* expression to be confined to a subset of DCLK1+ tuft cells (**Figure 5.1h**), mostly residing on the villus (Fig1i), which was also found by a recent report[193]. This could be a consequence of a gradual increase in *Chat* expression as tuft cells migrate along the crypt-villus axis, which is supported by higher GFP signal in cells on the villus compared to the crypt (**Figure 5.1j**).

To test this, we determined the localization of the three tuft cell subtypes from our *Chat*^{BAC}-eGFP scRNA-seq dataset. We selected the top marker genes of each cluster and plotted the center of mass of their expression on the crypt-villus axis, as determined by Manco *et al*.[15] (Fig1k). Additionally, we calculated single-cell signature scores along the zones of the crypt-villus axis, to further support regional abundance predictions for the three tuft types[15] (**Supplementary figure 5.1g**). This resulted in a zonated pattern, with the tuft-p cells predicted to be closest to the crypt, an intermediate pattern for tuft-1 cells, and the tuft-2 cells located highest up towards the villi tips (**Figure 5.1k** and **Supplementary figure 5.1g**). Considering the differential *Chat* expression in the three tuft cell subtypes (Fig1g), such zonated localization is in line with the increase in *Chat*^{BAC}-eGFP+ cells towards the villus, observed in our immunofluorescence experiments (Fig1hij). We further validated confinement of the proliferative tuft-p state to the crypt region by assessing EdU incorporation in DCLK1+ cells after a 4-hour pulse (**Figure 5.1l** and **Supplementary figure 5.1h**). These experiments also indicated that proliferation primarily precedes emergence of *Chat* expression (**Figure 5.1l**), in line with low *Chat* expression in tuft-p cells (**Figure 5.1g**).

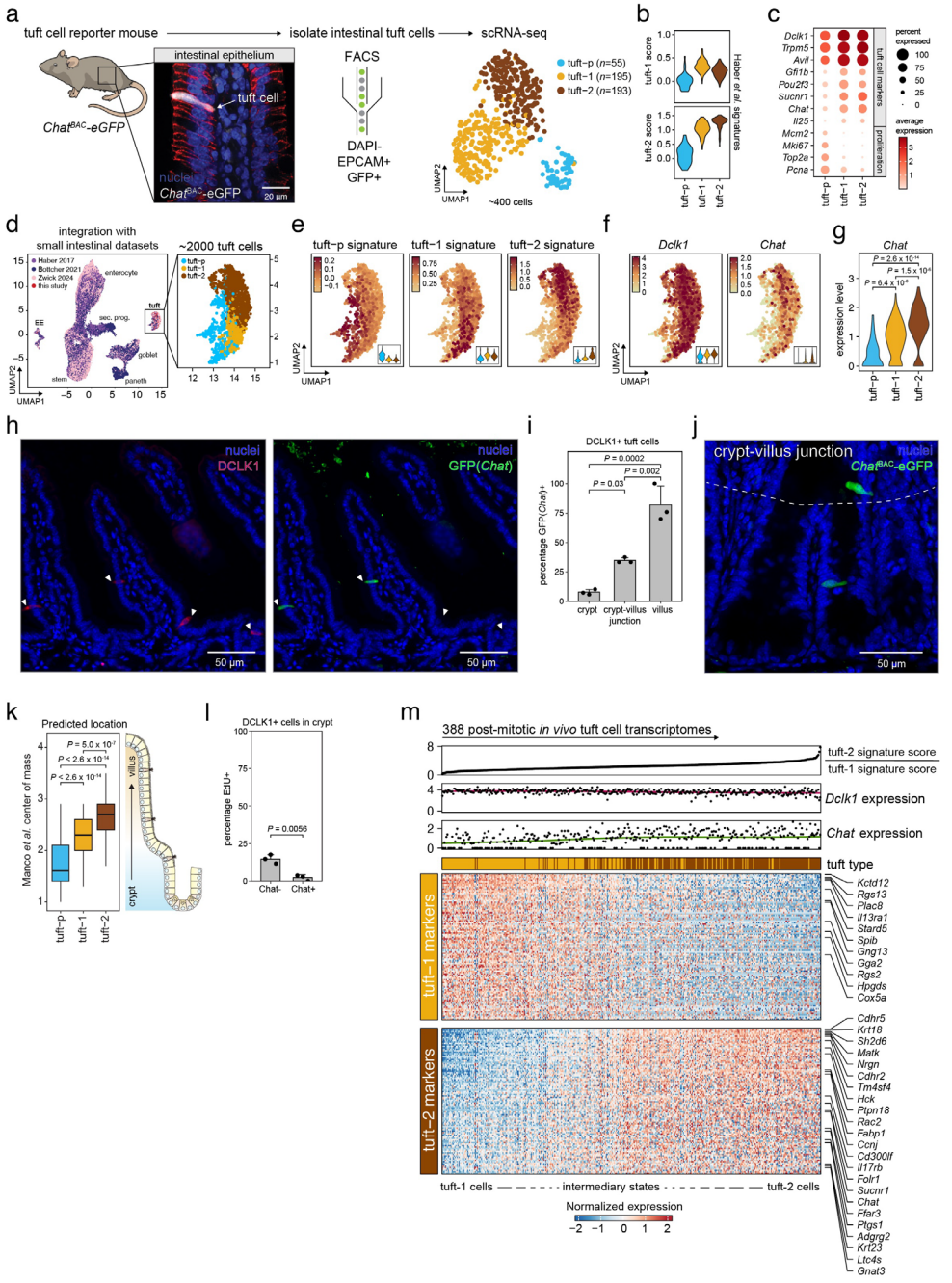


Figure 5.1: Gradual Chat expression marks a continuum of zonated tuft cell states on the crypt-villus axis. a) Left: zoom-in of intestinal villus where *Chat*^{BAC}-eGFP marks solitary tuft cells (white), with epithelial membrane (b-catenin, red) and nuclear staining (blue). Middle: FACS strategy for scRNA-seq of GFP+ cells from *Chat*^{BAC}-eGFP mice (n = 3). Right: UMAP with unsupervised clustering of 443 *Chat*^{BAC}-eGFP+ tuft cells. Tuft-p: tuft-precursors. **b)** Violin plots depicting distribution of tuft-1 and tuft-2 signature scores[17] per cluster. **c)** single-cell

dotplot showing expression of tuft cell and proliferation markers per cluster. **d)** Left: UMAP of integrated single-cell transcriptomic datasets of mouse small intestinal epithelium. Cells are colored by study origin. Right: zoom-in of tuft cell cluster (~2000 cells). EE: enteroendocrine cells. Sec. prog.: secretory progenitors. **e)** Relative expression of tuft-p, tuft-1 and tuft-2 signatures (top 100 differentially expressed genes) overlayed on UMAP of integrated tuft cell cluster and as violin plots (insets). Signatures were extracted from *Chat*^{BAC}-eGFP+ tuft clusters in panel a. **f)** Relative expression of *Dclk1* and *Chat* overlayed on UMAP of integrated tuft cell cluster and as violin plots (insets). **g)** Violin plot depicting distribution of relative *Chat* expression for all cells per cluster (ANOVA $p < 0.01$, Wilcoxon rank sum test).**h)** Fluorescent image of *Chat*^{BAC}-eGFP mouse small intestine stained for DCLK1 (red), GFP (green) and nuclei (blue). White arrows: DCLK1+ tuft cells. **i)** Quantification of frequency of *Chat*+ cells among DCLK1+ cells for indicated regions on the crypt-villus axis ($n = 3$ mice, 268 DCLK1+ tuft cells, ANOVA < 0.001, Tukey HSD test). **j)** Representative image showing lower expression of *Chat*^{BAC}-eGFP in tuft cells in crypt regions. White dashed line: crypt-villus junction. **k)** Zonation profile of tuft-p, tuft-1 and tuft-2 clusters (panel a). Boxplots show center of mass on the crypt-villus axis[15] for the top 100 cluster-specific genes (ANOVA $p < 0.0001$, Tukey HSD test). **l)** Barplot depicting frequency of EdU+ cells, split by Chat phenotype (Chat+ or Chat-), among DCLK1+ cells in the crypt ($n = 3$ mice, 112 DCLK1+ tuft cells, t-test). **m)** Heatmap showing co-expression of tuft-1 and tuft-2 marker genes (adjusted p-value < 0.01, Wilcoxon rank sum test). Cells from the post-mitotic tuft-1 and tuft-2 clusters (panel a) are ordered by the ratio of tuft-1 and 2 scores[17], indicated at the top of the heatmap. Relative *Dclk1* and *Chat* expression per cell are shown above the heatmap.

Taking into account the zonated localization of the post-mitotic tuft-1 and tuft-2 cells on the crypt-villus axis, we ranked cells from these clusters based on a tuft-2/tuft-1 signature score and visualized expression patterns of tuft-1 and tuft-2 marker genes in a heatmap (Fig1m). Notably, this revealed a gradual transition of tuft-1 to tuft-2 transcriptomic profiles via intermediary phenotypes encompassing mixed expression of both tuft-1 and tuft-2 markers. Crucially, this continuum of cell states and absence of binary confinement, suggests that the two previously reported tuft-types represent sequential differentiation states along the crypt-villus axes. As expected, *Chat* expression levels gradually increased from tuft-1 to tuft-2 states (Fig1m), in line with the predominant localization of *Chat*^{BAC}-eGFP+ cells towards the higher part of the villus (Fig1i), while *Dclk1* expression levels were high among all post-mitotic tuft cells (Fig1m). In sum, these descriptive *in vivo* data reveal early tuft precursors in the crypt and a continuum of tuft-1 and tuft-2 transcriptomic states along the crypt-villus axis, which are marked by gradually increasing *Chat* expression.

5.2.2 Small intestinal tuft cell phenotypes are conserved between mouse and human

To investigate if mature intestinal tuft cell phenotypes are conserved between mouse and human, we reanalyzed a recently published single-cell transcriptomic dataset of the human small intestinal epithelium, containing ~800 tuft cells[203] (**Figure 5.2a**). Unsupervised re-clustering of the tuft cells in the dataset resulted in three clusters, which we labeled tuft-p, tuft-1 and tuft-2 (**Figure 5.2b**) based on expression of the mouse tuft cell signatures (**Supplementary figure 5.2a**) and expression of proliferation markers (**Figure 5.2c; Supplementary figure 5.2b**). Recently, similar tuft subtypes were found in human organoids and intestinal tissue[204], including a proliferative state (tuft-3), akin to the tuft-p cells described here.

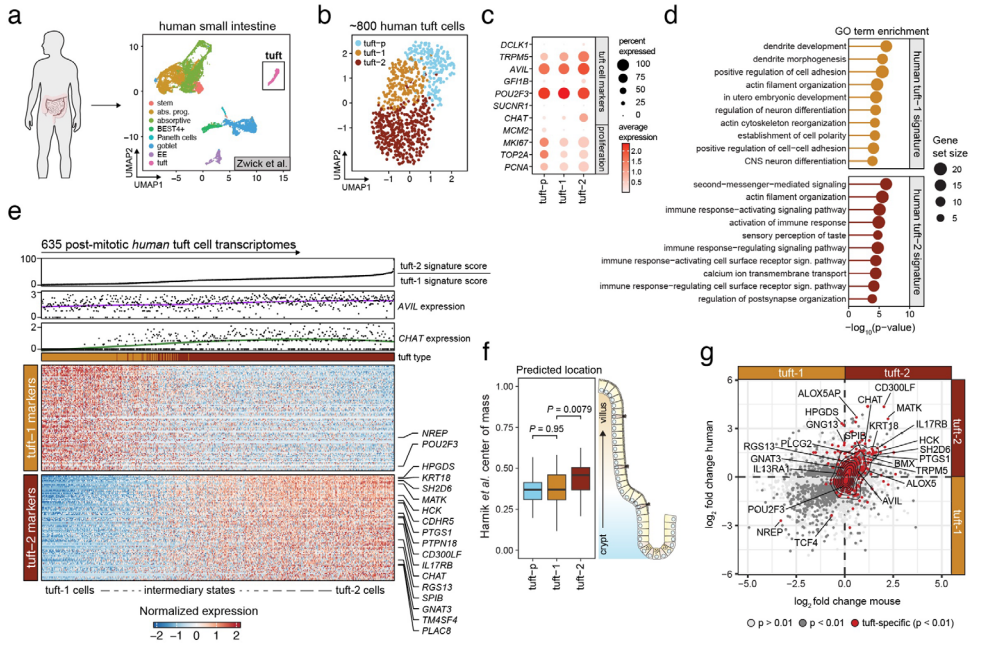


Figure 5.2: Small intestinal tuft cell phenotypes are conserved between mouse and human. a) Schematic and UMAP of published single-cell transcriptomics dataset of human small intestine[203]. Cell types are annotated by color. EE: enteroendocrine cells. Abs. prog.: absorptive progenitors. Sec. prog.: secretory progenitors. **b)** Unsupervised clustering of 844 human tuft cells[203] identifies three separate tuft cell populations. **c)** Single cell dotplot showing expression of tuft cell and proliferation markers per cluster. Average expression is represented by dot color while the percentage of expressing cells is denoted by the dot size. **d)** Lollipop plot showing GO-term enrichment analysis (clusterProfiler; Biological Process) of human tuft-1 and tuft-2 expression profiles. Dot size represents the number of genes within the gene set, stalk length corresponds to significance. **e)** Heatmap showing co-expression of tuft-1 and tuft-2 marker genes (adjusted p-value < 0.01, Wilcoxon rank sum test) in human post-mitotic tuft cells. Transcriptomic profiles are ordered by their ratio of human tuft-1 and 2 signature scores, indicated at the top of the heatmap. *AVIL* and *CHAT* expression per cell is shown above the heatmap. **f)** Zonation profile of human tuft-1 and tuft-2 populations from panel b. Center of mass on crypt-villus axis[205] for the top 50 cluster-specific genes is shown. Boxes represent interquartile range. Black bars represent median value (ANOVA p < 0.01, Tukey HSD test).

Next, we determined human tuft-1 and tuft-2 signatures and performed GO-term enrichment analysis (**Figure 5.2d**). Akin to mouse, this showed immune-related programs to be enriched in human tuft-2 cells, while human tuft-1 cells showed enrichment of neuron-related programs. Also in the human setting, *CHAT* was highest expressed in tuft-2 cells whilst pan-tuft cell markers like *AVIL* were uniformly expressed, and intermediary tuft phenotypes with cells expressing both tuft-1 and tuft-2 markers could be observed (**Figure 5.2e**). Moreover, previous spatial transcriptomics data[205] indicate that human tuft cells show a similar zonation profile along the crypt-villus axis as observed in mice, with tuft-2-specific markers found predominantly higher on the crypt-villus axis than tuft-1-specific markers (**Figure 5.2f**).

To perform a more comprehensive cross-species comparison of human and murine tuft cells, we established a core tuft expression profile shared between both species and determined

differentially expressed genes between the tuft-1 and tuft-2 clusters within each species (**Figure 5.2g**). Most tuft cell-specific genes were expressed in tuft-2 cells in both mouse and human, suggesting that the tuft-2 state is stronger conserved between mouse and human. Vice versa, most genes that show higher expression levels in tuft-1 cells are not unique to the tuft lineage. Taken together, these analyses show similar transcriptomic profiles for tuft cell subtypes in mice and human and suggest a conserved tuft cell differentiation trajectory in which tuft cell precursors transition through an early tuft-1 phenotype into a mature tuft-2 phenotype.

Differential expression analysis between tuft-1 and tuft-2 subtypes in mouse and human. Tones of gray represent significance level. Tuft-specific genes, shared between human and mice, are highlighted in red. Contour plot indicates density distribution of these shared tuft-specific genes.

5.2.3 Tuft subtype-specific reporter knock-ins enable identification of tuft-1 and tuft-2 transcriptional states *in vitro*

To experimentally investigate the differentiation trajectories that give rise to the continuum of tuft-1 and tuft-2 states and evaluate their chemosensory capacity, we designed tuft subtype-specific reporter organoids. This *in vitro* system enables functional investigation of rare tuft cells in a controlled environment and allows for on demand enrichment of tuft cell numbers and real-time imaging of their differentiation trajectories. We first established small intestinal organoids from *Chat*^{BAC}-eGFP mice, in which spontaneous differentiation towards GFP+ tuft cells was occasionally observed in the standard culture medium (**Figure 5.3a**). Notably, the fraction of GFP+ tuft cells barely increased in organoids after a commonly applied treatment with IL-4 and IL-13[192,206,207] (Fig3bc). Like *in vivo*, GFP+ cells in organoids represented a subset of DCLK1+ tuft cells, recapitulated preferential localization to the villus compartment (**Figure 5.3de**) and co-expressed mature tuft-2 markers, such as CD45 and FOLR1 (**Figure 5.1m**; **Supplementary figure 5.3**). In sum, these data suggest that *Chat*^{BAC}-eGFP in organoids, like *in vivo*, predominantly marks a subset of tuft cells that is enriched for mature tuft-2 markers.

Next, we used our integrated *in vivo* single-cell transcriptomic datasets of mouse small intestinal epithelium to identify more markers that can discriminate between tuft-1 and tuft-2 cells and are not expressed in other epithelial lineages (**Figure 5.3f**). We selected *Nrep* as a marker gene that demarcates tuft-1 transcriptomic states (**Figure 5.3fg; Supplementary figure 5.4ab**). Vice versa, we picked *Folr1*, which shows highest expression in mature tuft-2 states (Fig3fg; **Supplementary figure 5.4ab**). In addition, we chose *Gng13* as an exclusive tuft cell marker, which shows strong expression in tuft-1 and intermediary phenotypes (Fig3fg; **Supplementary figure 5.4ab**). Using RNAscope, we confirmed confined expression of the tuft-1 marker *Nrep* in the crypt, tuft-2 markers *Chat* and *Folr1* on the villus and *Gng13* along the entire crypt-villus axis (Fig3i). Differentially zonated co-expression of

combinations of these markers supported transitioning from tuft-1 to tuft-2 phenotypes on the villus (Fig3j; **Supplementary figure 5.4c**). Independent genetic knock-ins were integrated at the STOP-codon of these genes and contained a fluorescent mScarlet reporter (monomeric RFP variant) and diphtheria toxin receptor (DTR) that both become separated from the tuft cell markers during translation of the P2A peptides[208] (**Figure 5.3k** and **Supplementary Figure 5.4defg**). We generated all three reporter lines in the *Chat*^{BAC}-eGFP background, such that the eGFP signal can be leveraged as a reference to cross-compare the new tuft subtype-specific markers (**Figure 5.3l**). Last, we confirmed confinement of our markers to DCLK1+ tuft cells and zonated localization of tuft-1 and tuft-2 markers to the organoid crypt and villus compartments, respectively (**Supplementary figure 5.4hi**).

A prerequisite to study mature tuft cell functioning in organoids, is to confirm that tuft cell phenotypes are accurately recapitulated. To examine the degree of similarity between *in vivo* and *in vitro*-derived tuft cells, we sorted the various fluorescent cell populations from our organoids for scRNA-seq (**Figure 5.3m** and **Supplementary figure 5.5a**). The ensuing dataset contained >500 tuft cells (**Figure 5.3n**) with tuft-p, tuft-1 and tuft-2 phenotypes that closely resemble their *in vivo* counterparts as illustrated by strong enrichment of their respective signatures and marker genes (**Supplementary figure 5.5bcd**). Resemblance between *in vivo* and *in vitro*-derived tuft cells is further demonstrated by the many tuft cells that showed an intermediary phenotype with mixed expression of both tuft-1 and tuft-2 markers (**Figure 5.3o**). Cytokine treatment did not induce alternative tuft phenotypes that were not present in the standard culture medium, as illustrated by mixing of tuft cells derived from either medium on the UMAP and a lack of medium-specific tuft clusters (**Supplementary Figure 5.5e**). For a direct comparison, we plotted tuft-1 and tuft-2 signature scores from tuft cells originating from either the integrated *in vivo* dataset (**Figure 5.1d**) or our *in vitro* organoid dataset (**Figure 5.**3n). Despite a strong overlap, most *in vitro* tuft cells were representative of the tuft-1 phenotype (**Figure 5.3r**), presumably a consequence of the crypt-inspired signaling factors in the organoid medium.

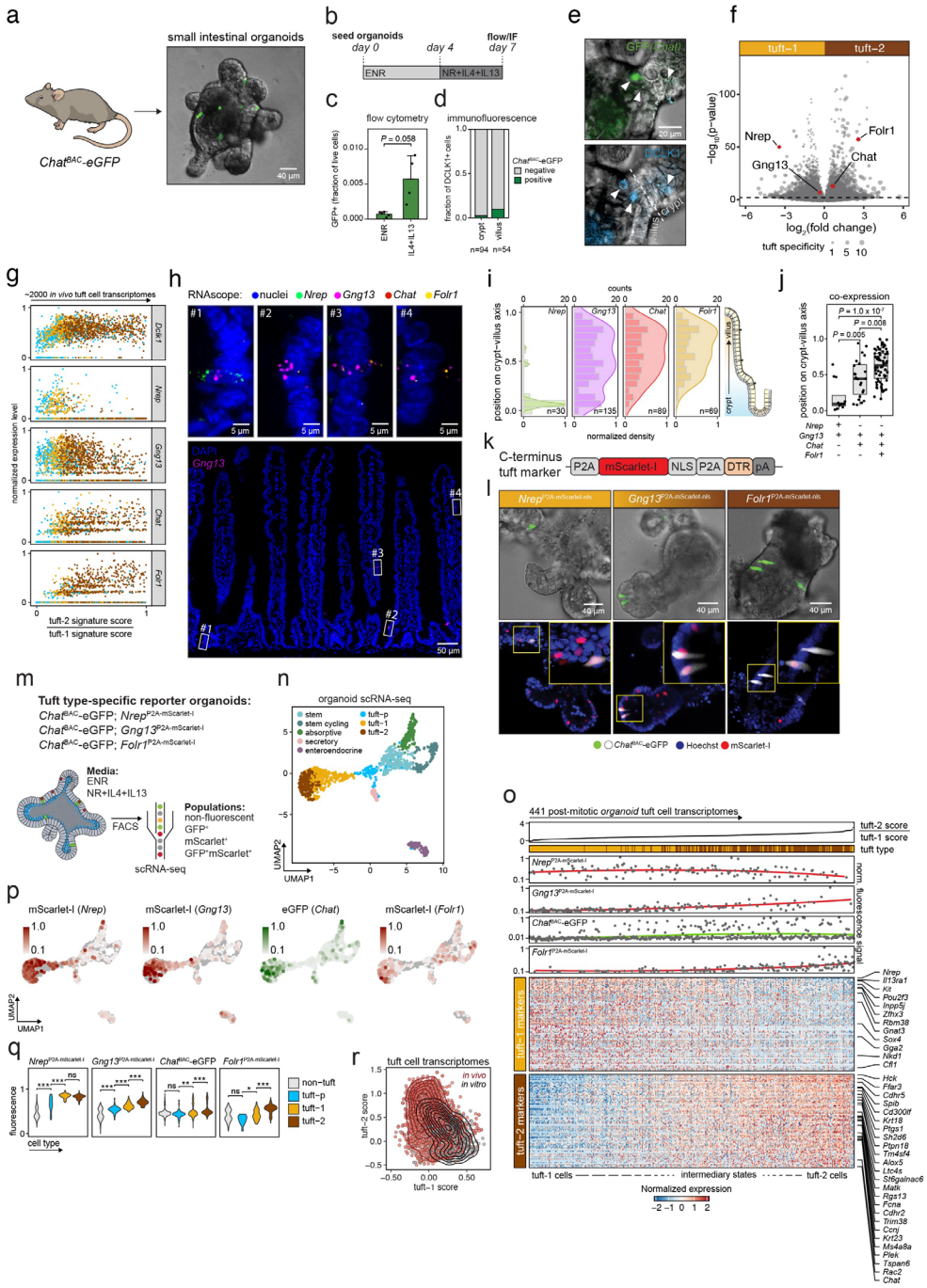


Figure 5.3: Tuft type-specific reporter knock-ins enable identification of tuft-1 and tuft-2 transcriptional states *in vitro*. **a)** Phase contrast image with fluorescent GFP (green) overlay of small intestinal organoid derived from *Chat*^{BAC}-eGFP mouse. **s)** Treatment regimen for tuft cell induction with IL-4 and IL13. E: EGF, N: Noggin, R; R-spondin, IF: immunofluorescence. ***c)*** Flow cytometry analysis of GFP(*Chat*)+ cells in organoids cultured in ENR

with or without IL-4 and IL-13 according to the regimen shown in b (t-test). E: EGF, N: Noggin, R; R-spondin. 5000 cells are shown per plot. **d)** Quantification of GFP(Chat)+ cells among DCLK1+ cells in crypt and villus region of organoids assessed by immunofluorescence following cytokine treatment as in b. **e)** Representative image of *Chat*^{BAC}-eGFP (top, green) organoid following cytokine treatment as in b and stained for DCLK1 (bottom, blue). White arrows: DCLK1+ tuft cells. White dashed line: crypt-villus junction. **f)** Vulcano plot showing differentially expressed genes (Wilcoxon rank sum test) between tuft-1 and tuft-2 clusters from the integrated mouse *in vivo* single-cell transcriptomic dataset (Figure 5.1d). Genes selected for reporter knock-ins are highlighted in red. Point size represents tuft specificity (see **Methods**). **g)** Relative expression levels of general tuft cell marker *Dclk1* and selected reporter genes in all tuft cells from the integrated mouse *in vivo* single-cell transcriptomics dataset (Fig1d). Cells are ranked based on their normalized tuft-2/1 signature score ratio. **h)** Representative fluorescence image of RNAscope FISH on mouse small intestinal tissue ($n = 4$) targeting *Nrep* (green), *Gng13* (magenta), *Chat* (red) and *Folr1* (yellow). Bottom: overview image showing entire crypt-villus units with nuclei (blue) and *Gng13* (magenta). Top: zoom-ins showing expression of the four markers at four positions along the crypt-villus axis. **i)** Location of cells expressing indicated tuft cell markers on crypt-villus axis. Numbers of analyzed cells are indicated. Histograms of counts and normalized density plots are shown. **j)** Co-expression of indicated tuft cell markers along the crypt-villus axis (ANOVA < 0.0001, Tukey HSD test). **k)** Schematic of knock-in template integrated at C-terminus of reporter genes. NLS: nuclear localization signal. mScarlet-I: monomeric red fluorescent protein. DTR: diphtheria toxin receptor. pA: polyadenylation signal. **l)** Top: Confocal brightfield images of tuft subtype-specific reporter knock-in organoids, overlayed with fluorescent *Chat*^{BAC}-eGFP signal (green). Bottom: fluorescent overlays of *Chat*^{BAC}-eGFP+ tuft cells (white) with indicated tuft-subtype reporters (red) and nuclear stain (blue). **m)** Experimental setup of reporter lines, culture conditions and flow cytometry to enrich for fluorescent tuft cells prior to plate-based single-cell transcriptomic analysis. **n)** UMAP with unsupervised clustering of organoid single-cell transcriptomic dataset, containing cells from all three knock-in tuft subtype reporter lines and both medium conditions. **o)** Heatmap showing co-expression of tuft-1 and tuft-2 marker genes (adjusted p-value < 0.01, Wilcoxon rank sum test) in organoid tuft cells. Organoid tuft cell transcriptomic profiles are ordered by the ratio of mouse tuft subtype signature scores (tuft-2 score / tuft-1 score). Original tuft subtype (yellow: tuft-1; brown: tuft-2) and normalized fluorescent mScarlet-I and eGFP signals of cells during sorting are indicated for each cell at the top of the heatmap. **p)** Normalized mScarlet-I and eGFP fluorescent signal intensities measured during cell sorting, superimposed on UMAP of organoid single-cell transcriptomic dataset. **q)** Violin plots of normalized fluorescent signal for each reporter per indicated cell cluster (ANOVA p < 0.001, ns not significant, **** adjusted p-value < 0.0001, * adjusted p-value < 0.05, Tukey HSD test). **r)** Contour plot showing density of tuft-1 and tuft-2 signature scores[17] ratios for tuft cells from the integrated *in vivo* dataset (Fig1d, red), as well as the organoid dataset (Fig3n, black), indicating underrepresentation of mature tuft-2 cells *in vitro*.

Next, we evaluated if our tuft cell subtype-specific reporters reliably indicate the intended tuft cell transcriptomic states. To this end, we superimposed the signal intensities of the fluorescent reporters, measured during cell sorting, on the single-cell transcriptomics data (**Figure 5.3opq**). GFP signal from the *ChAT*^{BAC}-eGFP was most dominant in cells with a high tuft-2 signature score (**Figure 5.3opq**). In contrast, for the tuft-1 marker *Nrep* we found high mScarlet signals in tuft-1 cells, which extended to cells with intermediary tuft phenotypes (**Figure 5.3opq**). As expected, *Gng13* shows a gradual increase of mScarlet signal starting in tuft-1 cells and becoming more abundant towards intermediary and tuft-2 expression profiles (**Figure 5.3opq**). *Folr1* marks cells with the highest tuft-2 signature scores (**Figure 5.3opq**). For all markers, we found strong concordance between fluorescent signal and gene expression (**Figure 5.3pqo; Supplementary Figure 5.5f**). Notably, despite being predicted as tuft-specific by the *in vivo* single-cell atlases (**Supplementary Figure 5.4b**) *Folr1* turned out to be less exclusive for the tuft lineage than the other markers (**Figure 5.3q; Supplementary Figure 5.5g**). Nonetheless, *Folr1*-driven mScarlet expression is a

reliable indicator of tuft-2 phenotypes within the tuft cell compartment and can therefore be used in combination with *Chat* to identify the most mature tuft cells (**Figure 5.3opq; Supplementary Figure 5.5g**). Together, these data show that tuft cell phenotypes, including intermediary states, are accurately recapitulated in organoids and they can be readily identified by our tuft type-specific reporters.

5.2.4 Cytokines drive the generation of new tuft-1 cells but are insufficient to advance their maturation

To study type 2 cytokine-induced tuft cell differentiation we characterized the frequency of tuft phenotypes in each of the three reporter lines under standard culture conditions (ENR) and upon tuft cell induction with the cytokines IL-4 and IL-13 (IL4+IL13) (**Figure 5.4a**). In agreement with our earlier observation that treatment with cytokines leads to a general increase in DCLK1+ tuft cell numbers but not *Chat*+ cells (**Figure 5.3bcd**), our knock-ins unequivocally demonstrate that the cytokines induce a strong increase in tuft-1 cells (**Figure 5.4bc**; mScarlet+ cells in the *Nrep* and *Gng13* reporter lines), but not tuft-2 cells (**Figure 5.4bc**; mScarlet+ cells in the *Folr1* reporter line and *Chat*^{BAC}-eGFP+ cells). Exposure to NOTCH inhibition, WNT stimulation, or longer treatment with cytokines, yielded similar skewed tuft subtype induction, with high frequency of tuft-1 cells and virtually no tuft-2 cells (**Supplementary Figure 5.6ab**). These observations are in stark contrast to type 2 cytokine-induced tuft cell hyperplasia *in vivo*, where both post-mitotic tuft cell subtypes increase in numbers, but most prominently the tuft-2 cells[17,207] (**Supplementary Figure 5.6cde**). Importantly, this reveals that cytokine administration to organoids, a commonly applied strategy to boost tuft cell numbers, hardly generates tuft-2 cells.

Since both the *in vivo* and *in vitro* single-cell transcriptome datasets indicate existence of intermediary tuft cell phenotypes, we treated organoids with cytokines to induce tuft cell specification and tracked the dynamics of tuft cell identity by means of fluorescent reporters, using confocal live-cell imaging (**Figure 5.4a**). First, we noticed clear differences in temporal expression patterns for the tuft-1 and tuft-2 markers within organoids (**Figure 5.4de**). Using *Chat*^{BAC}-eGFP as a reference signal, we observed that tuft-1 marker *Nrep* and tuft1/2 marker *Gng13* become active ahead of *Chat* expression, while the late tuft-2 marker *Folr1* becomes active in cells that were already positive for *Chat*^{BAC}-eGFP (**Figure 5.4de**). In agreement with the lack of binary confinement between tuft cell subtypes, these temporal expression patterns are indicative of a linear differentiation trajectory where tuft-1 cells transition into tuft-2 cells. This is further supported by the fact that tuft-1 marker expression predominantly emerged in the crypt compartment of organoids, while tuft-2 markers appeared in the villus compartment (**Figure 5.4f**).

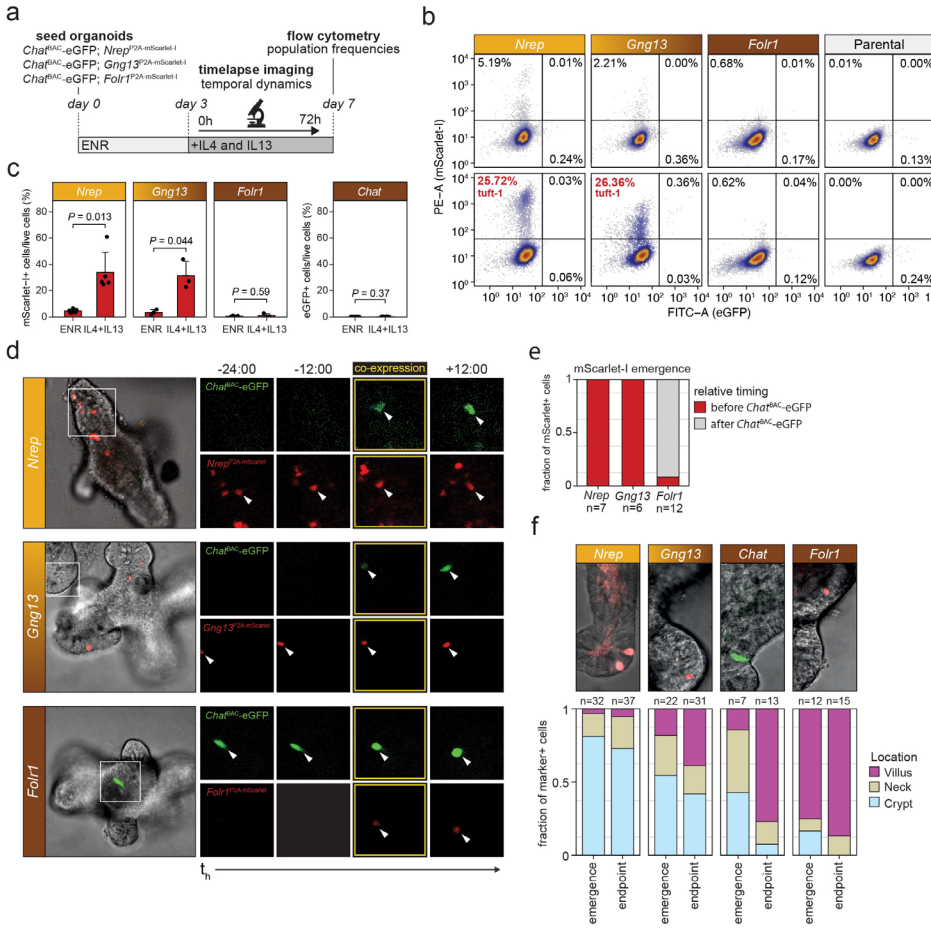


Figure 5.4: Cytokines drive the generation of new tuft-1 cells but are insufficient to advance their maturation. **a)** Schematic of experimental setup showing organoid knock-in reporters, treatment regimen and readout strategies (timelapse microscopy and flow cytometry). ENR: EGF, Noggin and R-spondin medium. **b)** Representative flow cytometry analysis of the three indicated tuft subtype-specific mScarlet- reporter lines and the parental *Chat*^{BAC}-eGFP line (parental). All lines were treated with IL-4 and IL-13 for 4 days, 3 days after seeding in ENR medium as illustrated in a. Percentages of the fluorescent populations are indicated in each plot. Individual panels contain combined data from 3 independent experiments, 5000 cells are shown per plot. **c)** Flow cytometry quantifications of tuft-subtype reporter frequencies (%) upon induction with indicated cytokines as in panel a (t-test). **d)** Microscopic stills from live imaging experiments demonstrating the temporal dynamics of tuft subtype reporters. Left: confocal brightfield image of organoids at the start of imaging, overlayed with fluorescent signals of indicated reporters. Right: stills of fluorescence channels (mSarlet-I: red, eGFP: green) from timepoints preceding and following co-expression of fluorescent markers to visualize differential dynamics with respect to *Chat*^{BAC}-eGFP. Cell that shows co-expression is indicated with a white arrowhead throughout imaging timepoints. **e)** Quantification of mScarlet-I emergence relative to *Chat*^{BAC}-eGFP expression for the three indicated tuft sutype-specific reporter knock-ins. **f)** Top: Representative confocal brightfield image with overlay of indicated fluorescent reporter. Bottom: Location distribution of tuft-reporter+ cells along the crypt-villus axes within organoids scored at emergence of reporter+ cells versus the end of imaging procedure.

Importantly, while type 2 cytokines elicit a massive increase in tuft-1 cells, this was not a consequence of tuft-1 proliferation, as we observed virtually no division of cells positive for our tuft-1 reporters in the time-lapse experiments. To confirm that our early tuft-1 marker *Nrep* marks post-mitotic tuft cells, we performed live imaging with fluorescent anti-CD24, a known tuft cell marker[193,209] that is already expressed in tuft-p cells (**Supplementary Figure 5.7a**). Indeed, anti-CD24 fluorescence was generally already detectable before *Nrep* expression emerged (**Supplementary Figure 5.7b**) and often marked dividing cells (**Supplementary Figure 5.7cde**). In contrast, rare dividing *Nrep*+ cells had typically initiated *Nrep* expression just prior to division (Supplementary **Figure 5.7de**). In conclusion, these experiments show that our tuft cell subtype-specific reporter knock-ins mark successive stages of maturation of post-mitotic cells committed to the tuft cell lineage. Moreover, *in vitro* treatment with IL-4 and IL-13 induces tuft cell specification to tuft-1 states, but is insufficient to efficiently advance their differentiation to the more mature tuft-2 cell states found on the villus.

5.2.5 Crypt-villus signaling gradients advance maturation from tuft-1 to tuft-2 states

Since tuft cell differentiation in cytokine-exposed organoids (IL-4 and IL-13) predominantly resulted in generation of tuft-1 phenotypes (*Nrep* and *Gng13* **Figure 5.5.4c**) but not tuft-2 phenotypes (*Chat* and *Folr1* **Figure 5.4c**), we set out to identify signals that may stimulate the transitioning of tuft-1 cells into mature tuft-2 cells. *In vivo*, tuft-2 markers are found predominantly on the villus (**Figure 5.1il** and **Figure 5.3ij**), which is reflected in our reporter organoids by a higher incidence of *Chat*+ and *Folr1*+ tuft-2 cells in the villus compartment of organoids (**Figure 5.4f,** >50 cells tracked). Therefore, we adjusted the cytokine differentiation regimen (**Figure 5.5a, #1**) to better represent a villus environment, starting with the depletion of stem cell-niche inspired growth factors *EGF, Noggin*, and *R-spondin* (**Figure 5.5a, #2**). Next, we screened for additional factors that are more abundant in the villus environment, such as *WNT5a*, *NRG1* and *BMP4*, and factors that are produced by tuft cells and could potentially act in an autocrine manner, such as *BMP2*, acetylcholine and *IL25* (**Figure 5.5a, #3**). Subtraction of the stem cell niche factors (ENR) from the medium resulted in higher frequencies of *Chat*^{BAC}-eGFP+ cells (**Figure 5.5b, #2**). Moreover, addition of villus and tuft cell produced factors further increased the *Chat*^{BAC}-eGFP+ cell population (**Figure 5.5b, #3**), with BMPs and acetylcholine eliciting the largest effects (**Supplementary Figure 5.8a**). Indeed, according to the single-cell transcriptome datasets, BMP receptors are expressed by all tuft cell subtypes and the BMP response genes *Id1* and *Id3* are higher expressed in tuft-2 cells (**Supplementary Figure 5.8b**), suggestive of a role for BMP signaling in differentiation to mature tuft-2 states. Concordantly, in *Bmpr1a* depleted intestinal epithelia[210] only the tuft-2, but not the tuft-1, transcriptomic signature is significantly reduced (**Supplementary Figure 5.8c**).

To test the effect of our villus-inspired maturation medium on the transitioning of tuft-1 to tuft-2 phenotypes, we evaluated their frequencies in our reporter organoids. Besides an overall increase in *Nrep*, *Gng13* and *Folr1* positive cells, this showed a clear shift in the maturation status of the tuft cell populations indicated by the emergence of double positive *Nrep*+*Chat*+ and *Gng13*+*Chat*+ tuft-2 cells as well as the most mature *Chat*+*Folr1*+ tuft-2 cells (**Figure 5.5cde**). Using flow cytometry time course experiments we observed these double positive mature tuft-2 phenotypes to arise late during the differentiation regimen (day 7), while an increase in *Nrep*+ tuft-1 cells was already observable two days prior (day 5) (**Supplementary Figure 5.9a**). With time-lapse imaging we observed , also in the villus-inspired maturation medium, *Chat*+ tuft-2 cells originating from *Nrep*+ tuft-1 cells (**Figure 5.5f**). Notably, switching to villus-inspired medium now resulted in decreasing *Nrep* expression (**Figure 5.5gh**), thus resembling more closely the transient *in vivo* expression, and was coupled with a late induction of *Folr1* expression (**Supplementary Figure 5.9b**)

To evaluate if the transient tuft-1 state is mandatory or facultative during tuft cell differentiation we used diphtheria toxin (DT) mediated subtype ablation via the DT receptor that was simultaneously knocked-in with the mScarlet reporters (**Figure 5.3k** and **Supplementary Figure 5.4d**). Depletion of early (*Nrep*+) tuft-1 cells with DT during our optimized differentiation regimen (**Figure 5.5i**), blocked the generation of *Chat*^{BAC}-eGFP+ cells (**Figure 5.5jk** and **Supplementary Figure 5.9c**). In contrast, under continuous depletion of late tuft-2 phenotypes (Folr1+), early tuft-2 *Chat*^{BAC}-eGFP+ cells did arise (**Figure 5.5jk** and **Supplementary Figure 5.9c**). These ablation experiments demonstrate that *Chat*+ tuft cells cannot arise from an alternative cellular source other than *Nrep*+ tuft-1 cells, in case of cytokine-induced tuft cell hyperplasia.

Lastly, we performed scRNA-seq on reporter organoids treated with our villus-inspired maturation medium to confirm a shift in maturation states and exclude induction of alternative non-physiological transcriptomic states. Unsupervised clustering with our previous organoid dataset showed mixing of tuft cells derived from the different medium conditions on the UMAP (**Supplementary Figure 5.10a**). Furthermore, tuft-1 and tuft-2 signature scores were similar in crypt and villus-inspired medium conditions (**Supplementary Figure 5.10b**). This indicates that, while the maturation medium had strong effects on the frequency of tuft-2 cells (**Figure 5.5cd**), it did not induce tuft cell transcriptomic states that were previously not observed. Moreover, trajectory inference through RNA velocity analysis corroborated the unidirectional differentiation trajectory from tuft-p to tuft-1 to tuft-2, which we repeatedly observed in our live imaging experiments (**Figure 5.5l**). Collectively, these experiments show that crypt-villus signaling gradients control the transitioning of tuft-1 to tuft-2 phenotypes (**Figure 5.5m**).

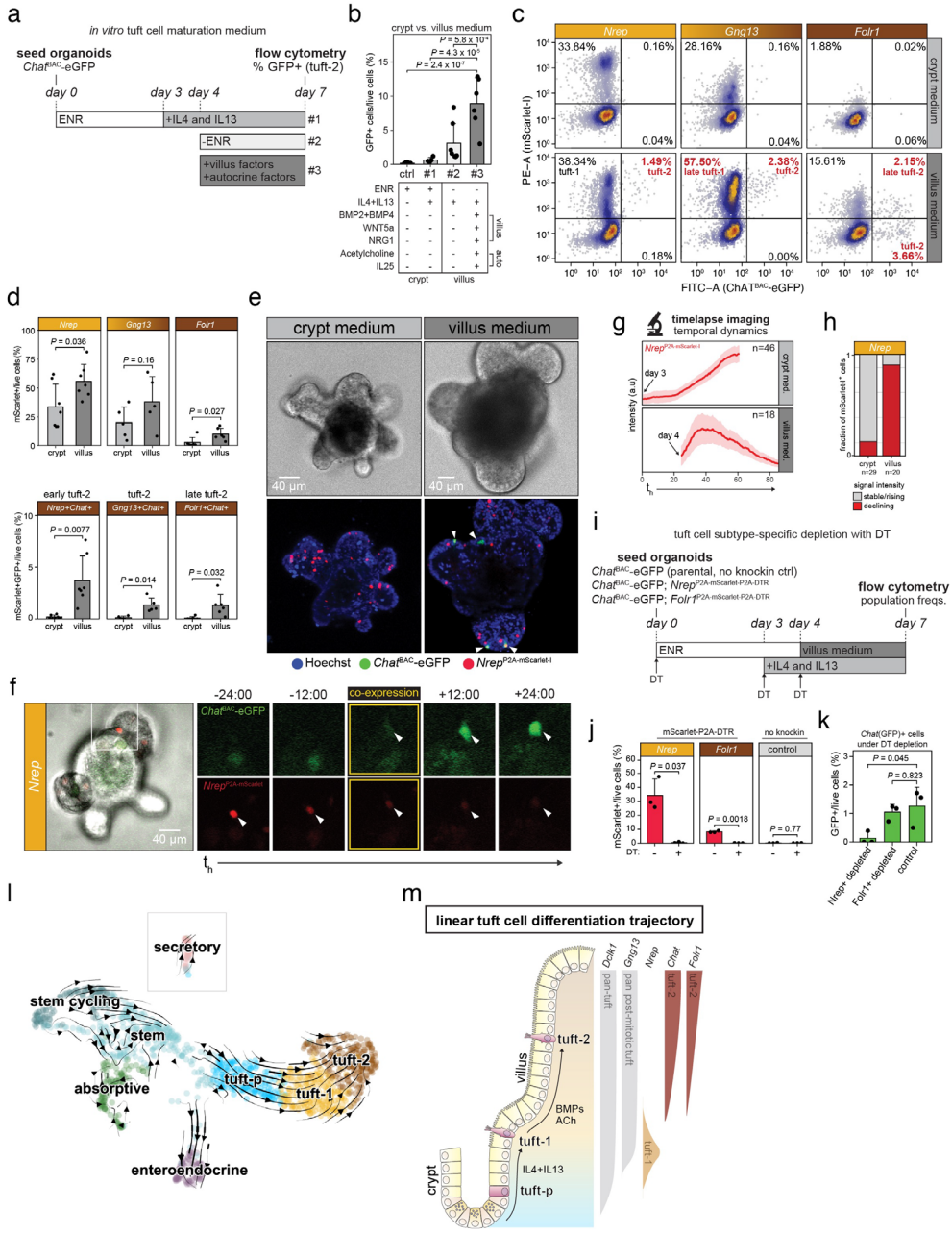


Figure 5.5: Crypt-villus signaling gradients advance maturation from tuft-1 to tuft-2 states. a) Experimental setup to promote tuft cell maturation with villus-inspired medium. *Chat*^{BAC}-eGFP reporter organoids were pretreated with IL-4 and IL13 on the 3rd day after seeding in ENR (EGF, R-spondin and Noggin) medium. On the 4th day, factors were added to, or depleted from, the medium. Frequency of GFP(*Chat*)+ tuft-2 induction was tested at day 7 with flow cytometry. **b)** Percentage of *Chat*^{BAC}-eGFP+ tuft-2 cells in organoids treated with IL-4 and IL-13 in presence or absence of crypt (#1) or villus/autocrine-inspired signaling factors (#2, #3) (ANOVA $p < 0.0001$, Tukey HSD test). **c)** Representative flow cytometry analysis of fluorescent population frequencies in indicated reporter

organoids treated with crypt (#1) or villus-inspired medium (#3) after IL-4 and IL-13 pretreatment. **d)** Percentage of mScarlet+ (top) and mScarlet+GFP+ (bottom) cells in indicated organoid reporter lines treated with crypt or villus-inspired medium (condition #1 vs #3, Figure 5.5ab) measured by flow cytometry (t-test). **e)** Representative fluorescence image of *Chat*^{BAC}-eGFP;*Nrep*$^{P2A\text{-}mScarlet\text{-}I}$ organoids treated with crypt or villus-inspired medium regimen (condition #1 vs #3, Figure 5.5ab). White arrows: *Chat*^{BAC}-eGFP+ cells. **f)** Experimental setup for tuft cell subtype-specific depletion with diphtheria toxin (DT). Organoids were treated as in condition #3 of Figure 5.5ab and DT was added with every medium change. **g)** Percentage of mScarlet+ cells in indicated reporter lines and parental no knock-in control (*Chat*^{BAC}-eGFP; no mScarlet-I or DTR) after differentiation during depletion with DT (t-test). **h)** Percentage of GFP+ cells in indicated reporter lines and parental no knock-in control (*Chat*^{BAC}-eGFP; no mScarlet-I or DTR) after differentiation during depletion with DT (ANOVA < 0.001, Tukey HSD test). **i)** Microscopic stills from live imaging experiments of *Chat*^{BAC}-eGFP;*Nrep*$^{P2A\text{-}mScarlet\text{-}I}$ organoids in villus-inspired medium (condition #3 of Figure 5.5ab). Left: confocal brightfield image of organoid at the start of imaging. Right: stills of fluorescence channels (mSarlet-I: red, eGFP: green) from timepoints preceding and following co-expression of fluorescent markers. Cell that shows co-expression is indicated with a white arrowhead throughout imaging timepoints. **j)** Average mScarlet-I signal over time in *Chat*^{BAC}-eGFP;*Nrep*$^{P2A\text{-}mScarlet\text{-}I}$ organoids treated with crypt or villus-inspired medium regimen (condition #1 vs #3, Figure 5.5ab). Shading in plot represents SEM. Number of cells comprising each graph are indicated. **k)** Barplot showing fraction of mScarlet-I+ cells in *Chat*^{BAC}-eGFP;*Nrep*$^{P2A\text{-}mScarlet\text{-}I}$ organoids with stable/rising or declining fluorescence signal in crypt or villus medium with cytokines (IL4+IL13). Related to j. Number of cells comprising each graph are indicated. **l)** RNA velocity-based trajectory inference with single cell transcriptomes from organoids treated with ENR, NR+IL4+IL13 or villus-inspired medium superimposed on UMAP. Direction of arrows predict unidirectional differentiation from tuft-1 to tuft-2 transcriptomic states. **m)** Schematic depicting linear model of intestinal tuft cell differentiation. Tuft-p: tuft precursors; IL4: interleukin-4; IL13: interleukin-13; ACh: acetylcholine; BMPs: bone morphogenetic proteins.

5.2.6 An organoid-based platform for functional characterization of tuft cell properties

Our tuft-type reporter organoids in combination with the optimized tuft maturation medium provide a powerful *in vitro* platform to study the basic cell biology that governs mature tuft cell functioning. To demonstrate its utility, we first set out to investigate the dynamics of the typical bottle-shaped tuft cell morphology[211] (**Figure 5.6a**), which features distinctive membrane characteristics. In addition to the well-known robust protrusions at the apical surface that form the tuft of the tuft cell, EM studies have also identified the presence of lateral microvilli as well as large basolateral extensions[212], which are speculated to be involved in cell-cell communication with neighboring cells[211,213-217]. Indeed, within our live-cell imaging recordings, we observed long dynamic protrusions. Subsequent imaging of *Chat*+ cells in combination with a genetically encoded nuclear marker (H2B-mScarlet) confirmed that these dynamic protrusions were basolateral extensions (**Figure 5.6b, Supplementary figure 5.11**), which presented with dynamic growth and shrink rates of more than 10 µm per hour (**Figure 5.6c**). Moreover, long protrusions (~20µm, up to 6 hours lifetime) showed more stability than short ones (5µm, one hour lifetime) (**Figure 5.6d**). Basal protrusions were also observed in *Nrep*+ tuft-1 cells (**Figure 5.6e**). These novel experimental opportunities highlight how future studies may improve our limited understanding of tuft cell morphological changes and its relation to functioning.

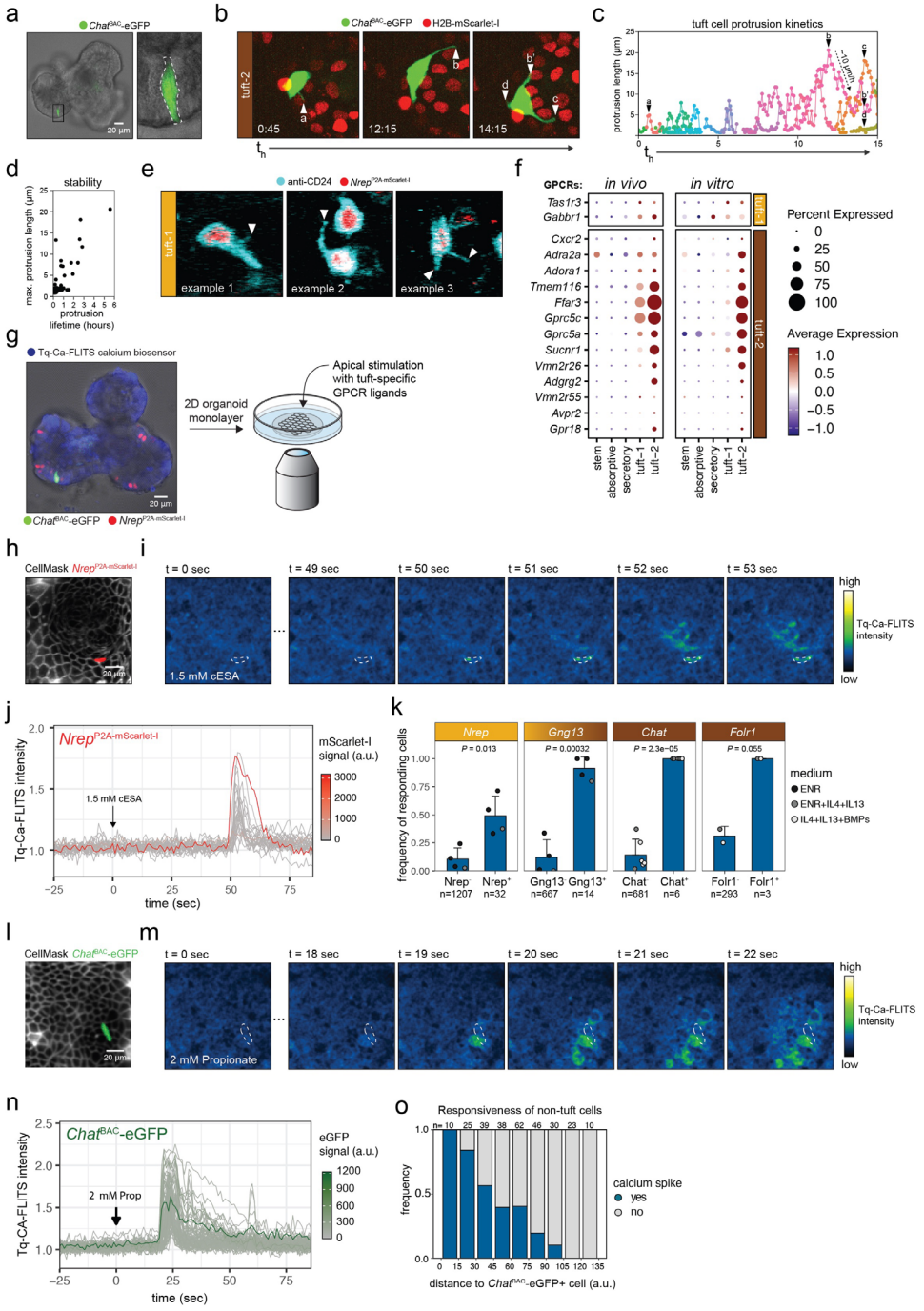


Figure 5.6: An organoid-based platform for functional characterization of tuft cell properties. a) Microscopic images of *Chat*^{BAC}-eGFP+ cell in organoid showing typical tuft cell morphology. Left: overview confocal brightfield image overlayed with GFP fluorescence channel. Right: zoom-in of indicated region. White

dotted line outlines *Chat*^{BAC}-eGFP+ tuft cell. **b)** Dynamic protrusions in *Chat*+ tuft-2 cells that can span several neighboring cells. Merge of fluorescence channels is shown: green, eGFP; red, nuclei visualized with H2B-mScarlet-I. Three stills of indicated timepoints are shown. White triangles with letters indicate protrusions. **c)** Protrusion lengths over time. Triangles with letters indicate single protrusion tracks corresponding to protrusions shown in microscopy images of b. **d)** Relationship between maximum protrusion length and protrusion lifetime. **e)** Microscopic stills from live imaging of *Nrep*$^{P2A\text{-}mScarlet\text{-}I\text{-}nls}$ organoids stained with anti-CD24, visualizing membrane protrusions of Nrep^{+} tuft-1 cells. Merge of fluorescence channels is shown: cyan, anti-CD24; red, *Nrep*$^{P2A\text{-}mScarlet\text{-}I\text{-}nls}$. White triangles with letters indicate protrusions. Three different examples of Nrep+ cells with dynamic protrusions are shown. **e)** Relative expression level of tuft-specific GPCRs within indicated cell types, as extracted from the mouse *in vivo* integrated single-cell dataset (Figure 5.1d) and organoids (Figure 5.3n). **f)** Experimental setup for live imaging experiments to monitor tuft cell activation. Tuft-subtype reporter organoids expressing the Tq-Ca-FLITS calcium biosensor are re-plated to form 2D monolayers that can be apically stimulated and are compatible with high temporal resolution imaging. **g)** Fluorescent image of 2D monolayer of reporter organoid with CellMask deep red incubation to label cell boundaries (membranes, white) and mScarlet-I reporter signal to label tuft identity. Image is prior to stimulation with cis-epoxysuccinic acid (cESA). **h)** Fluorescent intensity fluctuations of Tq-Ca-FLITS biosensor within the field-of-view cell monolayer as shown in panel d, following stimulation with cESA at t=0. Time (seconds) post cESA exposure is indicated at the top of each panel. White outline indicates Nrep^{+} tuft cell. **i)** Single-cell traces of Tq-Ca-FLITS intensity fluctuations within monolayer of panel e. Traces are colored by mScarlet-I signal measured in the corresponding cell. **j)** Responsiveness of tuft-reporter positive cells following stimulation with 1.5 mM cESA. Per bar, each point represents a separate experiment and is colored for the medium used (number of cells comprising each bar and p-values are indicated, t-test). **k)** As in panel d, but now monolayer of *Chat*^{BAC}-eGFP reporter organoid prior exposure to propionate. **l)** As in panel e, but Tq-Ca-FLITS intensity fluctuation post propionate exposure. Time since propionate addition is indicated at the top of each panel. Arrow indicates *Chat*^{BAC}-eGFP+ tuft cell. **m)** Single-cell traces of Tq-Ca-FLITS intensity fluctuations within monolayer of panel i. Traces are colored by GFP signal measured in the corresponding cell. **n)** Responsiveness of non-tuft cells (mScarlet-I^{-} GFP^{-}) in the proximity of the *Chat*^{BAC}-eGFP+ cell shown in h. Number of cells comprising each bar are indicated at the top of each bar. Results of one representative experiment are shown.

Next, we investigated tuft cell chemosensing that, similar to cells in the taste buds, involves GPCR-mediated signaling to control intracellular calcium pulses and subsequent membrane depolarization. Therefore, we complemented our tuft-subtype reporter organoids with a genetically encoded calcium biosensor (Tq-Ca-FLITS[218]), to visualize and measure the GPCR-mediated activation potential of putative stimuli (**Supplementary figure 5.12a**). We extracted tuft-specific GPCRs from our high-resolution integrated *in vivo* scRNA-seq dataset (**Figure 5.6f**). Reassuringly, the resulting list contained GPCRs that were previously found to be tuft-specific, such as *Sucnr1* and *Vmn2r26*, as well as some novel tuft-specific GPCRs, like *Gabbr1* and *Gpr18*. Most of the tuft-specific GPCRs recapitulated their *in vivo* expression pattern in organoids, with expression generally being highest in tuft-2 cells (**Figure 5.6f**). To develop a screening setup that enables apical stimulation and single-cell measurements of calcium influx in real-time, we converted calcium biosensor-expressing reporter organoids into 2D monolayers[219,220] [see Methods] (**Figure 5.6g**). We first exposed these monolayers to the succinate mimetic cis-epoxysuccinic acid (cESA), a known stimulus (**Figure 5.6h**). As expected, we observed calcium spikes within a minute after administration of the ligand (**Figure 5.6i**). While tuft cells were the first to respond to the stimulus (**Figure 5.6j**), calcium spikes did not remain exclusive to tuft cells and were similar in strength between tuft and non-tuft cells (**Supplementary figure 5.12b**). This propagation of calcium

pulses to nearby non-tuft cells is analogous to what has recently been observed in the trachea[221]. Concordant with the highest *Sucnr1* expression in tuft-2 cells (**Figure 5.6f**), we observed the highest response probability in cells positive for tuft-2 markers (**Figure 5.6k**). As a proof-of-principle of our functional platform to study chemosensing and screen for novel stimuli, we next tested responsiveness to the *FFAR3* ligand propionate, a microbiota-derived short chain fatty acid which is speculated but never formally shown to act directly on tuft cells. Stimulation with propionate readily triggered calcium spikes in mature *Chat*+ tuft cells (**Figure 5.6lm**), which propagated to nearby non-tuft cells (**Figure 5.6no**). Unlike stimulation with the succinate-mimetic cESA, *Nrep*+ tuft-1 cells showed no response to propionate (**Supplementary figure 5.12c**), underscoring the importance of studying chemosensing in fully mature functional tuft-2 cells.

5.3 Discussion

Tuft cells in the intestine function as epithelial orchestrators of host defense against luminal parasites[191-193,201]. Beyond their role during infections, alterations in intestinal tuft cell numbers and function are associated with immune-related diseases such as inflammatory bowel disease, coeliac disease and neoplasias[222]. Despite scientific and clinical interest in tuft cells, their rare occurrence *in vivo* and the absence of accurate *in vitro* models, have proven to be practical limitations that hamper investigation of the cell biological processes underlying tuft cell functioning. Here, we unlock organoids as an experimental model system to study tuft cell functioning *in vitro*. We show that the various transcriptomic states of tuft cells *in vivo* and their interdependencies are truthfully reflected in organoids, optimize culture conditions to enrich for the most mature tuft cell states, and showcase experimental opportunities to study the dynamics of distinctive tuft cell morphology, behavior and chemosensing properties at high resolution.

In our in-depth characterization of *in vivo* tuft cell phenotypes, upfront enrichment of *Chat*+ cells prior to single-cell transcriptomic profiling ensured a substantial number of high-quality tuft cell profiles. Through analysis of multiple scRNA-seq datasets from mouse, organoids and human, we consistently find three tuft cell subtypes marked by pan-tuft cell markers (e.g. *Dclk1* in mice and *POU2F3* and *AVIL* in mice and human). In addition to the earlier reported tuft-1 and tuft-2 transcriptomic states within the tuft cell pool[17], we also document precursor tuft cells (tuft-p) with proliferative capacity that resemble the recently documented proliferative tuft cells in human organoids (there called tuft-3)[204]. These tuft-p cells were already fully committed to tuft cell fate and typically exhibited at most one round of division, resembling similar precursor cells seen for other mouse intestinal secretory cells[88]. Moreover, our data revealed frequent presence of intermediate tuft cell profiles with co-expression of both tuft-1 and tuft-2 markers. This continuum of transcriptomic states, with no clear demarcation between the two previously described tuft-1 and tuft-2 types, argues against an early binary bifurcation to either of two mature tuft cell types during differentiation. Using

tuft subtype-specific reporter knock-ins in organoids, we investigated directionality over the spectrum of tuft profiles by real-time tracking of cell identity and showed that cytokine-induced tuft cell differentiation follows a linear trajectory from a tuft-p to a tuft-1 to a tuft-2 state. This indicates that the transcriptional programs in the tuft cell subtypes are expressed sequentially within the same cell and not by two distinct mature tuft cell subtypes and explains the earlier reported spatial preference of the tuft-1 and tuft-2 cells to the crypt and villus compartments, respectively[15]. Previous timelapse imaging of organoids from a dual *Trpm5*^{GFP}:*Adgrg2*^{mCherry} reporter mouse, in combination with histological analysis of intestinal tissues, also showed emerging expression of tuft-2 marker *Adgrg2* in cells that already expressed general tuft cell marker *Trpm5*, providing independent support for our observations[223].

Using our organoid reporters, we show that tuft cell maturation is stimulated by BMP signaling, which also supports functional zonation of enterocytes, goblet cells and enteroendocrine cells along the crypt-villus axis[14,210]. Previously, BMP2 signaling originating from tuft cells was shown to prevent *de novo* tuft cell specification from stem cells located at crypt bottoms[207]. This does not contradict our finding that BMP ligands, increasingly expressed by the intestinal stroma towards villus tips, promote maturation of already existing tuft cells. Similarly, microbiota-derived butyrate has been shown to limit tuft cell specification from stem cells[206], but its effect on tuft cells that populate villi is unknown.

While ChatBAC-eGFP was thought to label all cells committed to the tuft cell lineage within the small intestine[201], our transcriptomic datasets and organoid timelapse measurements indicate that *Chat* expression levels emerge in late tuft-1 cells and keep rising during tuft cell maturation. Even though we included cells with minor *Chat* expression levels in our analyses, our study mainly provides insights into the consecutive steps of tuft cell differentiation and maturation. In contrast, many studies have focused on the initial steps of stem cell specification towards the tuft lineage[195-199], and recently the capacity of tuft cells to act as reserve stem cells. While our data is less suited to investigate early fate decisions, we did observe a significant fraction of tuft cells with a progenitor phenotype. These tuft-p cells expressed pan-tuft cell markers, as well as markers of proliferation and other epithelial lineages, hinting at recent multilineage potential. Consequently, single marker genes specific to the tuft-p cells were not identified. This prohibited their genetic labeling with knock-in reporters in organoids and prevented investigations into how tuft-p cells are generated.

We found few phenotypic traits in tuft-1 cells that are unique to the tuft lineage, yet absent in tuft-2 cells. Rather they seem to express most tuft-specific genes at a lower level when compared to tuft-2 cells, akin to a more immature state. Of the few tuft-specific genes that showed higher expression in tuft-1 than in tuft-2, *Nrep* was the most prominent marker being exclusive to tuft-1 cells in both mice and human. Alternatively, most tuft-specific genes, and in particular GPCRs that play a central role in chemosensing, as well as immune effector molecules, showed increasing expression levels from tuft-1 to tuft-2 in both mouse and

human. Furthermore, our organoid models allowed us to show that both tuft-1 and tuft-2 cells generate highly dynamic basolateral protrusions of several tens of microns long. For future studies, it will be intriguing to investigate how diverse input signals are processed by tuft cells, how they are communicated to other cells, and the relationship between dynamic morphology and function.

Understanding the unique cell biology that underlies tuft cell functioning is of high interest to identify therapeutic targets to intervene with immune-related disorders and infections in the intestinal tract. The findings of our study show high similarity between the mature tuft cell phenotypes in mice, organoids and human. Moreover, mouse organoids are fully complementary with *in vivo* models and will be instrumental to generate mechanistic insights that can then be tested in physiological settings. Our comprehensive characterization and comparison of tuft cell phenotypes *in vivo* and in organoids, paired with reporter models and optimized differentiation strategies, established a functional platform to address the poorly understood mechanisms of perception, processing and transmission of chemosensory signals by tuft cells.

5.4 Methods

Animals. *Chat*^{BAC}-eGFP mice were obtained from the Jackson Laboratory and housed at the Hubrecht institute and Netherlands Cancer Institute mouse facilities under specific pathogen-free (SPF) conditions. For all experiments, adult mice at least 6 weeks of age, of either sex, were used. Animal experiments were approved by the IvD HI-KNAW and Netherlands Cancer Institute animal welfare committees and conducted according to relevant guidelines and regulations.

Visualization of *Chat*^{BAC}-eGFP in mouse small intestinal tissue. Slide preparation for single-cell-resolution imaging of fluorescent proteins in their three-dimensional near-native environment (Fig. 1aj) was performed as in *Snippert et al.*[224] In short, intestines were fixed in 4% paraformaldehyde (PFA) at room temperature for 20 min and washed in cold PBS. 1 cm^2 of intestinal wall was put in a mold. Four percent low melting point agarose (40°C) was added and allowed to cool on ice. Once solid, a vibrating microtome was used to make semi-thick sections (150 µm). Sections were stained with anti-mouse β-catenin (Sigma C2206, 1:200), overnight at 4 °C with mild rocking. After washing, sections were incubated with AF568 goat anti-rabbit secondary antibody (Invitrogen A11036, 1:500) and counterstained with Hoechst 33342 (ThermoFischer Scientific 62249, 1:2000). Sections were embedded in Vectashield (Vector Laboratories) after a final washing step. Image acquisition was performed on a Leica SP8 WLL scanning confocal microscope (40x-water-N.A.1.1) with Leica Application Suite X. Post-acquisition image processing and analysis were performed with FIJI ImageJ 2.16.0/1.54g.

EdU labeling and immunofluorescence of mouse small intestinal tissue. To label cells in S phase, 1 mg of EdU (200 µl in PBS) was injected intraperitoneally 4 hours prior to euthanasia. Mouse intestines were excised, rinsed with PBS, opened longitudinally, and fixed overnight in 1% PFA. Following fixation, intestines were washed with PBS, rolled into Swiss rolls, and embedded in parrafin for formalin-fixed parrafin embedded (FFPE) tissue blocks or OCT compound (frozen at -80 °C) for cryosectioning. FFPE samples (Fig1h) were sectioned and slides were stained with anti-DCLK1 (Abcam ab37994, 1:1000) and anti-GFP (Aves GFP-1020, 1:4000). OCT-embedded tissues (**Figure 5.1il**) were cryosectioned and stained with anti-DCLK1 (Abcam EPR6085, 1:1000) and anti-GFP (Abcam EPR14104, 1:1000). EdU-positive cells were visualized using the Click-iT reaction, according to the manufacturer's instructions (Click-iT EdU Imaging Kit, Thermo Fisher Scientific/Invitrogen C10340). Secondary antibodies included: Donkey anti-goat IgG Alexa Fluor 488 (A-11055, 1:1000), Donkey anti-rabbit IgG (H+L) Alexa Fluor 568 (A-10042, 1:1000) and Alexa Fluor 647 from the Click-iT reaction (C10340, 1:200). DAPI (1:000) was used for nuclear counterstaining. Cryosections were mounted using antifade mounting medium (Vectashield, Vector Laboratories). Imaging was performed with an Olympus VS200 slide scanner (FFPE) or a Leica TCS SP8 confocal microscope (cryosections; 25x water immersion objective; HC

FLUOTAR L, N.A. 0.95 W, VISIR, 0.17 mm cover glass, FWD 2.4 mm). Post-acquisition analysis and quantification were performed with QuPath 0.5.1.

RNAscope fluorescent *in situ* hybridization assay. Prior to FISH, small intestinal tissues were fixed and embedded as described above to make FFPE tissue blocks. FFPE blocks from *Chat*^{BAC}-eGFP mice or wildtype C57BL/6 mice (acknowledgements) were cut (5 µm slice thickness) and stained with probes targeting *Nrep* (Biotechne 562221), *Gng13* (Biotechne 462531-C4), *Chat* (Biotechne 408731-C3) and *Folr1* (Biotechne 575401-C2) using the RNAscope Multiplex Fluorescent Detection Kit v2 (Biotechne 323110) according to the manufacturer's instructions. In brief, paraffin sections were deparaffinized, treated with hydrogen peroxide for 10 min and boiled in target retrieval buffer for 15 min before a 30-min protease treatment. Probes were amplified and detected using fluorescent probes based on opal dyes. Slides were counterstained with DAPI for 30s, mounted using ProLong Gold Antifade Mountant (Thermo Fisher scientific) and imaged on an Olympus VS200 slide scanner. Post-acquisition analysis and quantification were performed with QuPath 0.5.1.

Organoid derivation and culture. Small intestinal crypts were isolated as described previously[11]. In brief, mouse small intestines were excised, rinsed with PBS and incubated with DPBS containing 5 mM EDTA for 40 min at 4 °C on a carousel to extract crypts. Crypts were pelleted by centrifugation of the supernatant, washed with advanced DMEM (Gibco) and plated in drops of Matrigel (Corning) to form organoids. Organoids were derived and cultured with ENR medium: advanced DMEM/F12 (Gibco) supplemented with 1% Penicillin/Streptomycin (P/S, Lonza), 1% HEPES buffer (Gibco) and 1% GlutaMAX (Gibco), 5% R-spondin conditioned medium (in-house production), 10% Noggin conditioned medium (in-house production), 1x B27 (Invitrogen), 1.25 mM N-acetylcysteine (Sigma-Aldrich) and 50 ng/ml EGF (Invitrogen). Organoids were maintained at 37 °C with 5% CO_2 and passaged weekly through fragmentation by shear-stress (pipetting). Medium was refreshed every 2 days.

Organoid tuft cell differentiation. 3D organoid cultures were subjected to differentiation regimens as indicated. In general, tuft cell specification was induced on the 3rd day after passaging with ENR medium supplemented with recombinant murine IL-4 and IL-13 (10 ng/ml each, Immunotools). Thereafter, tuft cell maturation was facilitated by removal of EGF, Noggin and R-spondin from the medium and addition of one of the following or a combination thereof on the 4th day after passaging: recombinant murine IL-25 (20 ng/ml, Immunotools), recombinant human BMP2 (20 ng/ml, Immunotools), recombinant human BMP4 (20 ng/ml, Immunotools), recombinant human NRG1 (20 ng/ml, R&D systems), recombinant murine WNT5a (20 ng/ml, Biotechne) and acetylcholine chloride (100 µM, Sigma Aldrich). For investigation of alternative tuft cell inducing factors (**Supplementary Figure 5.. 6ab**), DAPT (5µM Sigma D5942) and Wnt-surrogate FC fusion protein (0.5nM, U-Protein Express) were used. For DT depletion experiments, medium was supplemented

with 50 ng/ml Diphtheria toxin (Sigma D0564). Organoids were harvested for downstream analysis on the 7th day after passaging unless otherwise stated.

Tissue and organoid dissociation for flow cytometry and sorting. For isolation of tuft cells from mouse tissue, small intestines were excised, rinsed with PBS and incubated for 10 min in PBS containing 100 μM DTT. Next, the small intestines were cut into smaller fragments, transferred into DPBS containing 5mM EDTA and spun on a carousel for 40 min at 4 °C. Organoids were extracted from Matrigel with cold advanced DMEM. To make single-cell suspensions, tissue fragments or organoids were trypsinized for 5 min at 37 °C with TypLE containing 10 μM Y-27632 Rho-kinase inhibitor to inhibit anoikis. Trypsinized single-cell preparations were washed with advanced DMEM, filtered (40 μm cell strainer) and resuspended in advanced DMEM containing 10 μM Y-27632. In case of immunofluorescence antibody stainings, single-cell suspensions were stained for 30 min on ice with APC anti-mouse EpCAM (Biolegend 118213, 1:100) and AF700 anti-mouse CD45 (Biolegend 103127, 1:100). Cells were stained briefly with DAPI prior to flow cytometry.

Flow cytometry. Single-cell suspensions were prepared as described above. Flow cytometry measurements were performed on a BD FACSCelesta CellAnalyzer. Single live cells (DAPI negative) were gated and GFP and mScarlet-I fluorescence was measured in the FITC-A and PE-A channels, respectively. Gates were set based on negative control samples. Flow cytometry data was analyzed and visualized using BD FACSdiva software 9.2 and the flowCore (2.14.0) and CytoExploreR (1.1.0) R packages.

Plate-based scRNA-seq. Single-cell suspensions were prepared as described above. Viable single cells (DAPI negative) were sorted (BD FACSAria III) into 384-well cell-capture plates from Single Cell Discoveries, which contain a 50 nl droplet of well-specific barcoded primers and 10 μl of mineral oil (Sigma M8410). After sorting, plates were briefly centrifuged at 500 × *g* and then kept on dry ice till further storage at –80 °C. Single-cell RNA sequencing was performed by Single Cell Discoveries according to an adapted version of the SORT-seq protocol[225] with primers described in van den Brink *et al.*[226]. Cells were heat-lysed at 65 °C followed by cDNA synthesis. After second-strand cDNA synthesis, all the barcoded material from one plate was pooled into one library and amplified using in vitro transcription (IVT). Following amplification, library preparation was performed following the CEL-Seq2 protocol[227] to prepare a cDNA library for sequencing using TruSeq small RNA primers (Illumina). The DNA library was sequenced by paired-end sequencing on an Illumina Nextseq™ 500, high output, with a 1×75 bp Illumina kit (read 1: 26 cycles, index read: 6 cycles, read 2: 60 cycles).

scRNA-seq analysis. For alignment of reads, an adopted version of the nf-core scrnaseq pipeline (2.4.0)[228] was used (https://github.com/gowanaka/nf-core-scrnaseq). In brief, STARsolo (2.7.10b) was used to align reads to a custom mm10 transcriptome including *eGFP* and *P2A-mScarlet-I-P2A-DTR* transgenes and ERCC spike-ins. Following mapping, count matrices were generated with STARsolo (2.7.10b). Gene expression was analyzed

using Seurat (5.0.1)[229]. Cells with <30% mitochondrial content, <25% exogenous ERCC spike-in content and >1000 detected genes were selected for downstream analysis. Mitochondrial transcript counts were removed prior to count normalization and scaling by the Seurat NormalizeData and ScaleData functions, respectively. Seurat CCA-integration was performed to eliminate technical plate-based batch effects and subsequently unsupervised clustering was used to cluster cells according to the standard Seurat workflow. Gene expression signature scores were calculated with the Seurat AddModuleScore function. Differential expression analysis was performed with the FindAllMarkers function. Tuft specificity scores for Figure 5.3f were calculated as $\log_{10}((1/\text{average expression in non-tuft epithelial cells})+1)$. Integration with published single-cell transcriptomic datasets (Haber *et al.* GSE92332, Zwick *et al.* GSE201859 and Böttcher *et al.* GSE152325) was performed with the Seurat package (harmony integration) according to the standard workflow. Clusters were annotated according to the cell type annotations included with the published datasets. GO-term enrichment analysis was performed with clusterProfiler (4.8.3)[230]. RNA velocity analysis was performed with scVelo (0.3.2)[231].

Analysis of published RNA sequencing data. Transcript counts from mice infected with *Nippostrongylus brasiliensis* and *Bmpr1a* knockout mice were obtained from E-MTAB-9183[207] and GSE194004[210], respectively. The R package DESeq2 (1.40.2)[232] was used to perform differential expression analysis, followed by gene set enrichment analysis (GSEA) with clusterProfiler (4.8.3)[230] and visualization with enrichplot (1.20.3).

Generation of organoid knock-ins. Organoid knock-ins were generated by in-trans paired nicking (ITPN) as described in Bollen and Hageman *et al*[208] (Supplementary figure 5.4defg). Cas9 D10A nickase (addgene #48141) locus-specific expression vectors were generated according to published protocols[233]. For transfection, organoids were trypsinized to cell clumps containing ~5 cells (1×10^6 cells total) and coelectroporated with 15 µg DNA and 5 ug of Cas9 D10A nickase and targeting vector using the NEPA21 Super Electroporator (Nepagene) following described conditions[234]. Electroporated cell clumps were plated in Matrigel overlayed with ENR culture medium, which was supplemented with 10 µM Y-27632 Rho-kinase inhibitor and 0.25 nM Wnt surrogate-FC fusion protein (U-Protein Express) for the first 3 days. Targeted cells were selected using 2 µg/ml puromycin and maintained as polyclonal populations. Proper integration of the knock-in constructs was validated by targeted PCR of genomic organoid DNA (isolated with QIAamp DNA Micro Kit from Qiagen, according to the manufacturer's instructions), followed by Sanger sequencing of the PCR product using the following primers:

Nrep untargeted allele	fw	5' CTGCTGGGTTCTGAGGGAGGCT
	rv	5' GTGACTGCACATCCACCAGGGC
Nrep 5' knock-in junction	fw	5' CTGCTGGGTTCTGAGGGAGGCT
	rv	5' CTCGCCCTTGCTCACAGG
Nrep 3' knock-in junction	fw	5' GCAACCTCCCCTTCTACGAG

	rv	5′ GTGACTGCACATCCACCAGGGC
Gng13 untargeted allele	fw	5′ CACTGGTTCCTTGCGAGCCCTG
	rv	5′ TCCCAAGGAGCCACCTGTTGCT
Gng13 5′ knock-in junction	fw	5′ CACTGGTTCCTTGCGAGCCCTG
	rv	5′ CTTCAGCTTGGCGGTCTG
Gng13 3′ knock-in junction	fw	5′ GCAACCTCCCCTTCTACGAG
	rv	5′ TCCCAAGGAGCCACCTGTTGCT
Folr1 untargeted allele	fw	5′ CAGTGGTGGGAGGACTGCCAGA
	rv	5′ GCCCCTCCCCGCAAACAAAACT
Folr1 5′ knock-in junction	fw	5′ CAGTGGTGGGAGGACTGCCAGA
	rv	5′ CTTCAGCTTGGCGGTCTG
Folr1 3′ knock-in junction	fw	5′ GCAACCTCCCCTTCTACGAG
	rv	5′ GCCCCTCCCCGCAAACAAAACT

Lentiviral transduction of organoids. Organoids were transduced with lentivirus encoding hEF1a-Tq-Ca-FLITS or H2B-mScarlet-I followed by an IRES and a blasticidin resistance cassette. In brief, organoids were incubated in trypsin at 37 °C to make a suspension of cell clumps containing approximately 5-10 cells. Cell clumps were transduced by spinoculation for 1h at room temperature and then plated in Matrigel overlayed with ENR culture medium, which was supplemented with 10 μM Y-27632 Rho-kinase inhibitor and 0.25 nM Wnt surrogate-FC fusion protein (U-Protein Express) for the first 3 days. Transduced organoids were selected using 3 μg/ml blasticidin (InvivoGen) and maintained as polyclonal populations.

Calcium imaging of 2D organoid monolayers. For calcium imaging of 2D monolayers, organoids were fragmented and seeded on polyacrylamide gels (18 kPa stiffness) that had been coated overnight with 100 μg/ml Laminin (Sigma-Aldrich L2020) and 250 μg/ml Collagen I (First Link (UK)) as in Pérez-González *et al*.[219] After seeding, organoid fragments were overlayed with ENR medium (described above), which was supplemented with 1x N2 Supplement (Thermo Scientific)**,** 10 ng/ml human-FGF2 (Peprotech), 3 μM Chiron (Bio-Connect)**,** 10 μM Y-27632 and 100 ug/ml primocin (InvivoGen), and contained a higher amount of R-spondin conditioned medium (20%). Chiron and Y-27632 were removed from the medium one day later. For differentiation of 2D monolayers, IL-4 and IL-13 (10 ng/ml each) were added 2 days before imaging, followed by addition of BMP2 and BMP4 (20 ng/ml each) in combination with removal of Noggin from the medium on the day prior to imaging. Timelapse imaging was performed 3-4 days after seeding of organoid crypts, when monolayers formed mature crypt-villus compartments, on a Leica SP8 WLL scanning confocal microscope (40x-water-N.A.1.1, 37 °C, 6% CO_2) with Leica Application Suite X at 1 frame per second using resonant scanning (8000 Hz). The genetically encoded Tq-Ca-FLITS(van der Linden et al., 2021) calcium biosensor was excited with an ultraviolet laser at

405 nm, to minimize bleed through of EGFP in the Tq-Ca-FLITS channel, and emission was measured at 445-505 nm. Organoid monolayers were stimulated during imaging with cis-epoxysuccinic acid (cESA, 1.5 mM, Fischer Scientific) or propionate (2 mM, Sigma-Aldrich P1880), as indicated. CellMask Deep Red (1:20000, Invitrogen, 640 nm excitation and 655-710 nm emission) was added 30 min prior to imaging to stain plasma membranes and enable post-acquisition cell segmentation. Before acquisition of each timelapse, EGFP (488 nm excitation and 495-545 nm emission) and mScarlet-I (570 nm excitation and 590-630 nm emission) signals were measured to determine tuft reporter status per cell in the field of view.

Live organoid timelapse imaging. For live imaging, organoids were mechanically fragmented and seeded in cold BME (Trevigen) in a four-well chambered cover glass (#1.5 high-performance cover glass, Cellvis). After seeding, wells were placed on a cold block for 10 minutes allowing the organoids to sink to a position close to the glass, as described earlier[88]. Organoids were cultured for 2 to 3 days in ENR medium until clear crypt-villus structures were apparent. Next, IL4 and IL13 (10 ng/ml each) were added to the medium and organoids were imaged for 3 days, starting on the 1st or 5th day after interleukin treatment to capture transitioning fluorescence phenotypes of all tuft reporters. For live staining of CD24, APC anti-CD24 (ThermoFischer 17-0242-80) was used. Imaging was performed at 37 °C and 5% CO_2 with a scanning confocal microscope, using either the Leica TCS SP8 with a 40× water immersion objective (numerical aperture, 1.10) or the Nikon A1R MP with a 40x oil immersion objective (NA = 1.30). The voxel size used was 0.4x0.4x2 micron for both microscopes. The maximum time resolution used was 24 minutes per frame.

Timelapse analysis. All tracking and timelapse analyses were performed using OrganoidTracker 2.0[53,235]. To measure marker intensities the average fluorescent intensity around the nucleus center was measured using a Gaussian kernel with a width of 1.5 pixels. To avoid measuring bleed through from the green (*Chat*^{BAC}-eGFP) channel we subtracted the green channel from the red channel (all other markers) such that there was no signal in the cytoplasm, where Nrep/Gng13/Folr1$^{P2A\text{-}mScarlet\text{-}I\text{-}nls}$ marker signal should be absent. Tuft cell locations were determined by evaluation of transmission signal. Cells in the spherical end of organoid buds were categorized as located in the crypt, while cells that fitted into a sphere drawn around the central villus region were deemed to be in the villus. Cells in the narrow region between crypt and villus compartment were deemed to be in the neck region. The crypt and the neck region were combined when analyzing the colocalization of ChAT and DCLK1 (**Figure 5.3d**). For the analysis of the relative timing of marker expression and the locations of emergence, data from at least three replicates was pooled. For the *Nrep*$^{P2A\text{-}mScarlet\text{-}I}$ signal in villus-inspired and standard crypt medium, a representative organoid was analyzed.

Automated tuft cell protrusion analysis. To track protrusions, we first extracted a 3D cell mask by thresholding on the *Chat*^{BAC}-eGFP signal. We then performed binary erosion and dilation operations to extract the mask of the cell body. The difference between the cell body and the full cell mask gave us the protrusions. Protrusion tips were defined by the pixel in

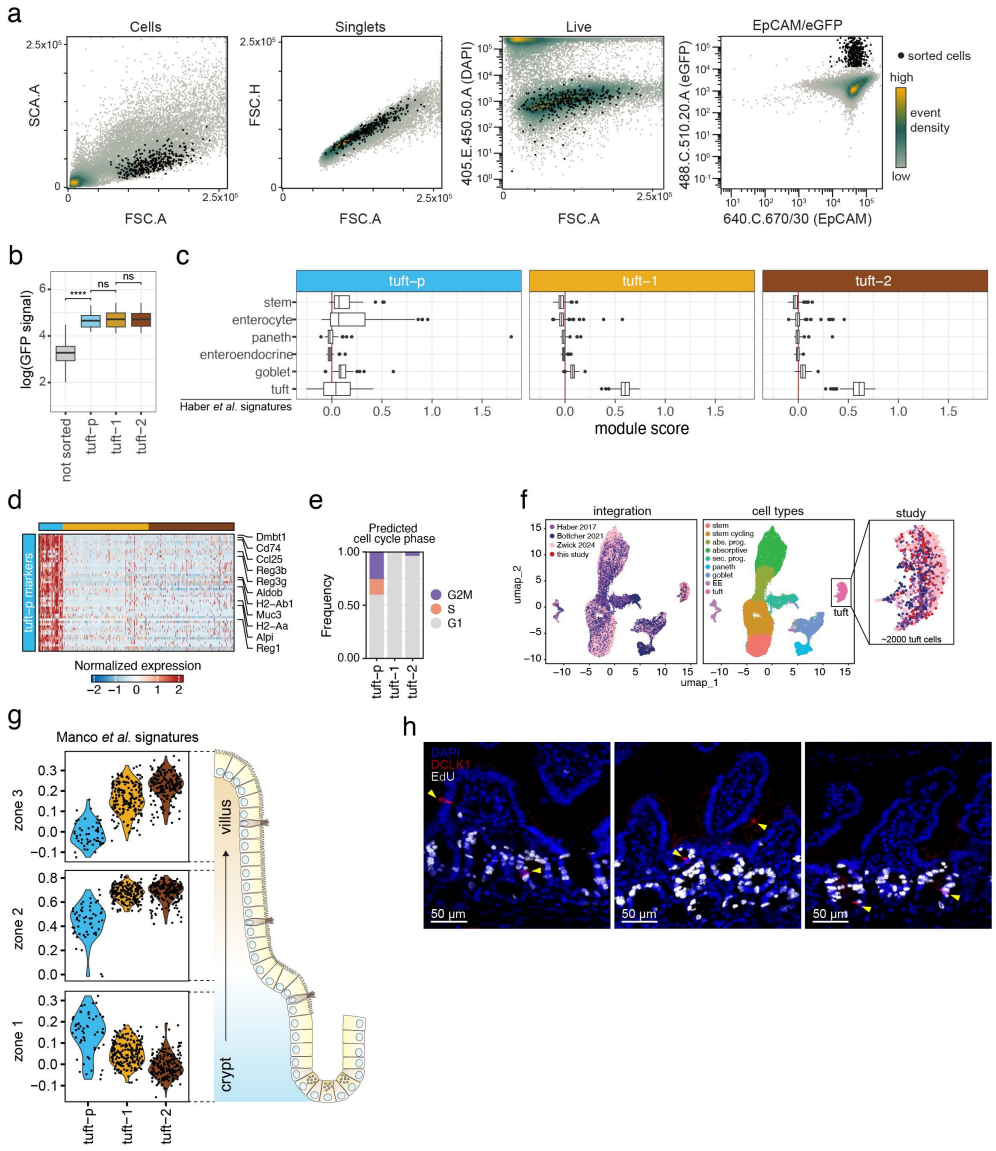
the protrusion that was furthest away from the cell body. Protrusions were linked by connecting protrusions to the protrusion with which they shared maximum overlap in the next frame. If there was little or zero overlap (for instance if protrusions were very thin or small), protrusions were connected with the tips closest to each other. Using the nuclear marker, tuft cells were automatically rotated such that the epithelial plane was in XY. For this, we locally fitted an ellipsoid to the nuclear signal, so that the two major axes defined a plane tangential to the epithelial layer. These could then be used to find the proper rotation needed.

Immunostaining of organoids. Immunofluorescence of organoids was performed as in Dekkers *et al.*[236]. Briefly, organoids were harvested from Matrigel droplets with cold advanced DMEM/F12 (Gibco) supplemented with 1% Penicillin/Streptomycin (P/S, Lonza), 1% HEPES buffer (Gibco) and 1% GlutaMAX (Gibco) and pelleted by centrifugation. Following fixation with 4% paraformaldehyde in PBS at 4 °C for 45 min, organoids were permeabilized with 0.1% (vol/vol) PBS-Tween and blocked with PBS containing 0.2% bovine serum albumin and 0.1% Triton X-100 (4 °C for 15 min). Primary antibodies (anti-mouse CD45, 1:200, Biolegend 103102; anti-mouse Folr1, 1:100, R&D Systems AF6936-SP) were incubated overnight in blocking buffer at 4 °C with mild rocking. After washing in blocking buffer (4x), secondary antibodies (AF568 donkey anti-Sheep, Invitrogen A21099, 1:500; AF568 goat anti-rat, Invitrogen A11077, 1:500) were incubated, together with Hoechst 33342 (ThermoFischer Scientific 62249, 1:2000), in the same manner. Finally, organoids were washed in blocking buffer (4x) and mounted on glass slides in a fructose-glycerol clearing solution (60% (vol/vol) glycerol and 2.5 M fructose). Image acquisition was performed on a Leica SP8 WLL scanning confocal microscope (40x-water-N.A.1.1) with Leica Application Suite X. Post-acquisition image processing and analysis were performed with FIJI ImageJ 2.16.0/1.54g.

Quantification and statistics. Statistical analysis was performed as noted in figure legends using R (R base, ggplot2 (3.5.1), ggpubr (0.6.0) and Seurat packages (5.0.1)). All experiments were performed in multiple distinct replicates. Data distribution was assumed to be normal, but this was not formally tested. Statistical tests were two-tailed Student's t-tests. For comparisons between more than two sample groups, one-way ANOVA was performed, using Tukey HSD for post-hoc analysis. Data are presented as mean +/- SD, unless otherwise stated in the figure legend. For boxplots, boxes represent interquartile range with central lines marking median value. For single-cell dotplots, average expression is represented by dot color and percentage of expressing cells is denoted by the dot size.

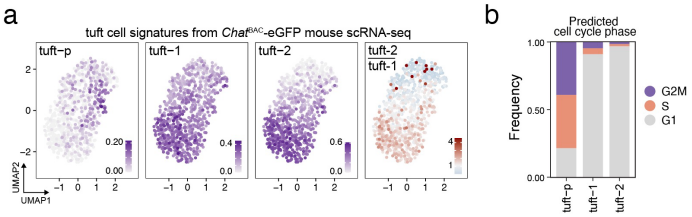
Data availability. Organoid and primary tissue single-cell RNA sequencing data presented in this study will be made publicly available upon publication and can be accessed at GSE276950. Source data are provided with this paper.

5.5 Supplementary figures

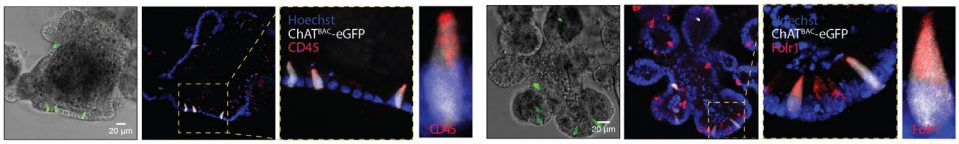


Supplementary figure 5.1: scRNA-seq of mouse small intestinal *Chat*^{BAC}-eGFP+ tuft cells. Related to Figure 5.1. **a)** Gating strategy used during FACS prior to plate-based scRNA-sequencing of EpCAM+*Chat*^{BAC}-eGFP+ tuft cells isolated from mouse small intestinal tissue. Event density is indicated by color gradient. Sorted events are denoted in black. **b)** Mean fluorescence signal (eGFP) of cells from indicated UMAP clusters measured by FACS prior to plate-based scRNA-sequencing (ns not significant, **** p<0.001, t-test). **c)** Box plot indicating the scores by which various small intestinal cell type signature(Haber et al., 2017) can be detected in single cells of the indicated tuft clusters as identified in the *Chat*^{BAC}-eGFP dataset. **d)** Gene expression profile of tuft-p cells. The top 50 differentially expressed genes (adjusted p-value < 0.01, Wilcoxon rank sum test) with the highest fold-change are plotted and selected genes are highlighted. **e)** Distribution of computationally predicted cell-cycle phase per cluster. **f)** UMAP of integrated single-cell transcriptomic datasets of mouse small intestinal epithelium. Cells are

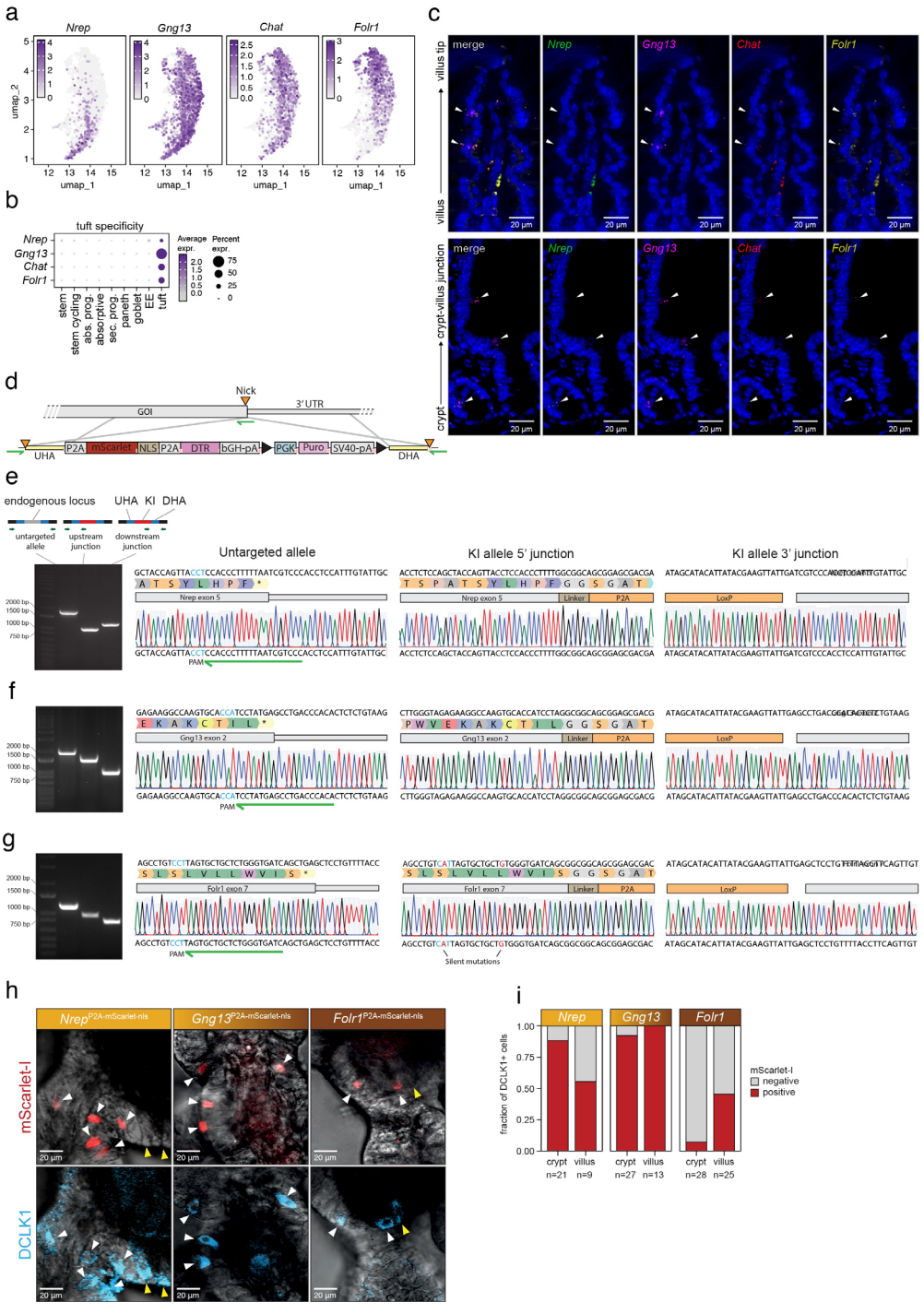
colored by study origin (left) or cell type (middle). Right: zoom-in of tuft cell cluster (~2000 cells) colored by study origin. EE: enteroendocrine cells. Abs. prog.: absorptive progenitors, Sec. prog.: secretory progenitors. **g)** Violin plots indicating the scores by which the zonated expression signatures along the crypt-villus axes(Manco et al., 2021) are detected in single cells of the three tuft clusters found in the *Chat*^{BAC}-eGFP dataset. **h)** Representative fluorescence microscopy images of EdU staining on mouse small intestinal tissue after 4-hour EdU pulse. Yellow arrows indicate DCLK1+ cells.



Supplementary Figure 5.2: Analysis of transcriptomic heterogeneity within the human small intestinal tuft cell pool. Related to Figure 5.2. **a)** UMAP of re-clustered human tuft cells overlayed with tuft-p, tuft-1 and tuft-2 gene signatures extracted from the *Chat*^{BAC}-eGFP mouse scRNA-seq dataset from figure 5.1. The right most panel shows the ratio of tuft-2 over tuft-1 signature expression. **b)** Distribution of computationally predicted cell-cycle phase per cluster.

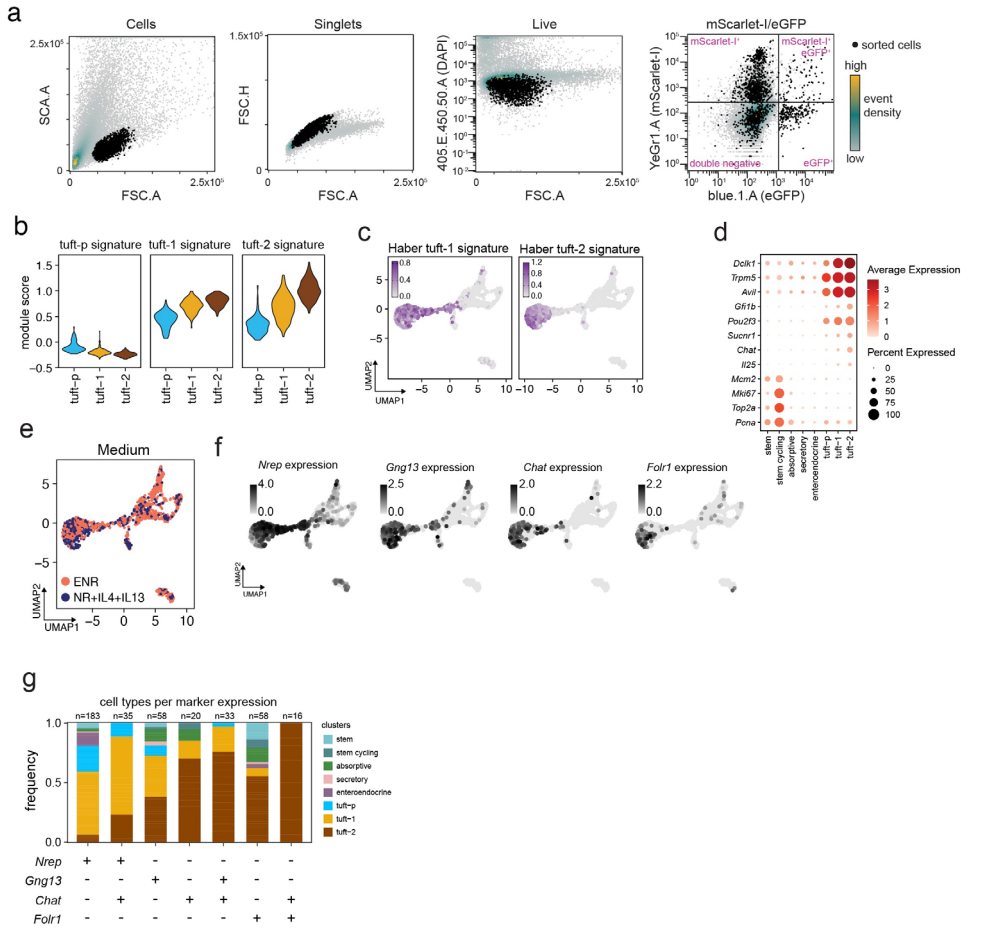


Supplementary Figure 5.3: Mature tuft cell markers are expressed in mouse small intestinal organoids. Related to Figure 5.3. Microscopy pictures of mouse small intestinal organoids from *Chat*^{BAC}-eGFP mice showing brightfield image (grey) overlayed with fluorescent GFP signal (green). From each organoid the corresponding fluorescent images are shown in three magnifications (GFP signal in white) with immunofluorescence staining against indicated tuft cell markers CD45 or FOLR1 (red). For counterstaining, nuclei (blue) are labelled with Hoechst 3342.

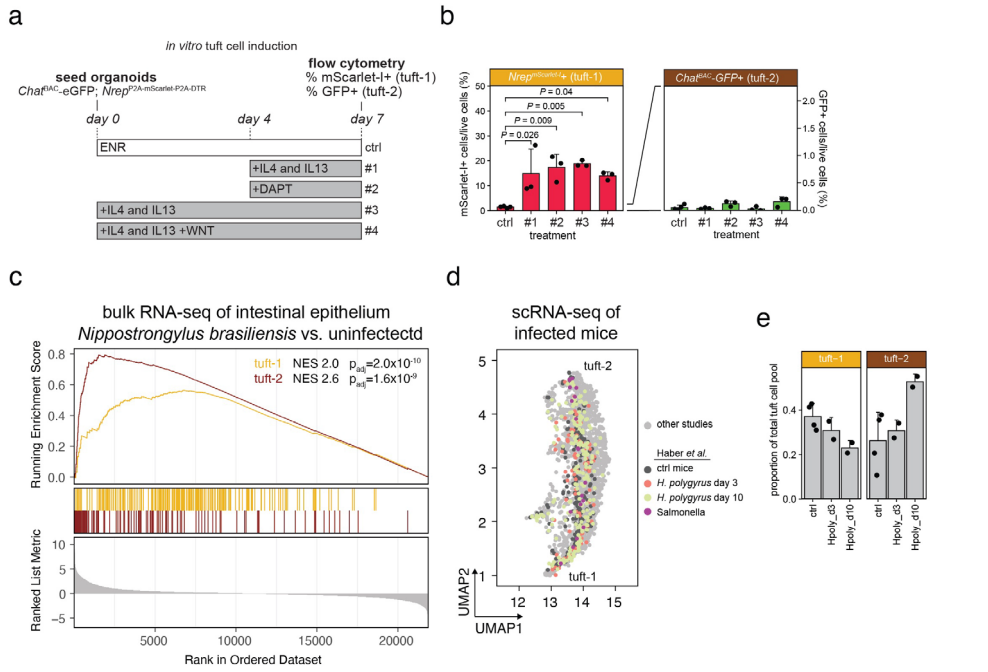


Supplementary Figure 5.4: Generation of tuft subtype-specific reporter knock-ins in mouse small intestinal organoids. Related to Figure 5.3. **a)** UMAP of integrated *in vivo* tuft cell cluster (**Fig1d**), overlayed with heatmap

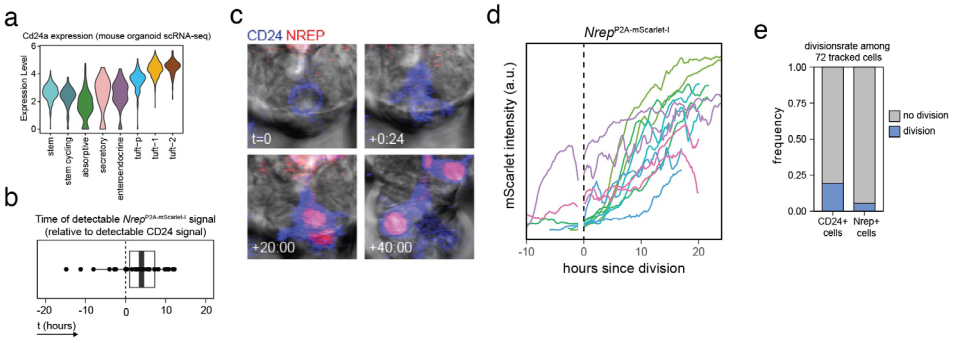
showing relative expression of selected reporter genes for prospective tuft type-specific knock-ins. **b)** Relative expression level of selected tuft subtype-specific reporter genes in various intestinal cell types featured in the integrated *in vivo* mouse small intestine single-cell transcriptome dataset (Fig1d). Average expression is represented by dot color. Percentage of expressing cells is denoted by dot size. EE: enteroendocrine, abs. prog.: absorptive progenitor, sec.prog.: secretory progenitor. **c)** Representative fluorescence microscopy images of RNAscope targeting indicated prospective tuft subtype markers. White arrows indicate cells positive for at least one of these markers. **d)** Schematic representation of genetic knock-in strategy using in-trans paired nicking, where Cas9D10A nickase generates single strand breaks ('nick', orange triangles) in genomic target site, as well as at the extremities of the homology arms in the knock-in donor template. Guide RNA (green arrows) targets vicinity of C-terminal STOP-codon. Black triangles indicate LoxP sites. UTR: untranslated region, UHA: upstream homology arm, DHA: downstream homology arm, NLS: nuclear localization signal, DTR: diphtheria toxin receptor, bGH-pA: polyadenylation signal of bGH, PGK: PGK promotor, Puro: puromycin resistance cassette and SV40-pA: polyadenylation signal of SV40. **e)** Validation of correct genomic integration of heterozygous knock-in in *Nrep* by PCR (left) and Sanger sequencing of the corresponding PCR amplicons. Untargeted allele remains unaffected with in-trans paired nicking. UHA: upstream homology arm, DHA: downstream homology arm, KI: integrated knock-in, PAM: protospacer adjacent motif. Guide RNAs are indicated with green arrows. **f)** As in d, but for *Gng13*. **g)** As in d, but for *Folr1*. **h)** Representative brightfield microscopy images overlayed with fluorescence of indicated tuft cell subtype reporter knock-in lines (mScarlet-I, red) stained with anti-DCLK1 (blue). White arrows indicate co-expression, yellow arrows indicate DCLK1+ cells without co-expression of tuft subtype marker. **i)** Quantification of h.



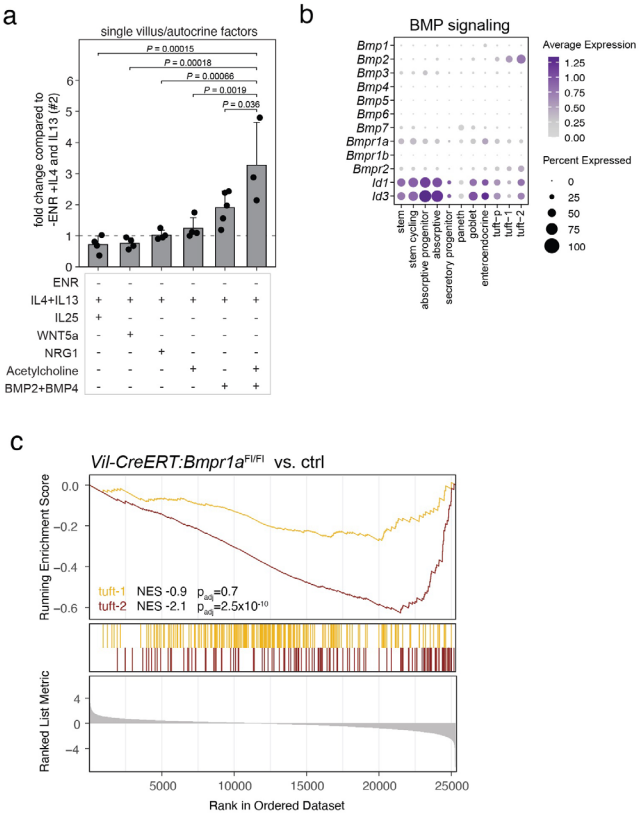
Supplementary Figure 5.5: scRNA-seq of tuft subtype-specific reporter knock-in organoids. Related to Figure 5.3. **a)** Gating strategy used during FACS prior to plate-based scRNA-sequencing of fluorescent cell populations from tuft type-specific reporter organoids. Event density is indicated by color gradient. Sorted events are denoted in black. Plot includes subsampling of events from all three organoid reporter lines and both medium conditions. **b)** Violin plot depicting tuft-p, tuft-1 and tuft-2 signature expression in *in vitro* organoid scRNA-seq tuft cell clusters. **c)** UMAP of organoid single-cell transcriptomic dataset overlayed with relative tuft-1 and tuft-2 signature scores[17]. **d)** Single cell dotplot showing expression of tuft cell and proliferation markers per tuft cell cluster in *in vitro* organoid scRNA-seq dataset. Average expression is represented by dot color while the percentage of expressing cells is denoted by the dot size. **e)** UMAP of organoid single-cell transcriptomic dataset, colored by medium condition. **f)** UMAP of organoid single-cell transcriptomic dataset, overlayed with normalized RNA expression of indicated tuft cell subtype markers. **g)** Barplot showing cell types from organoid single-cell transcriptomic dataset positive for indicated tuft subtype markers.



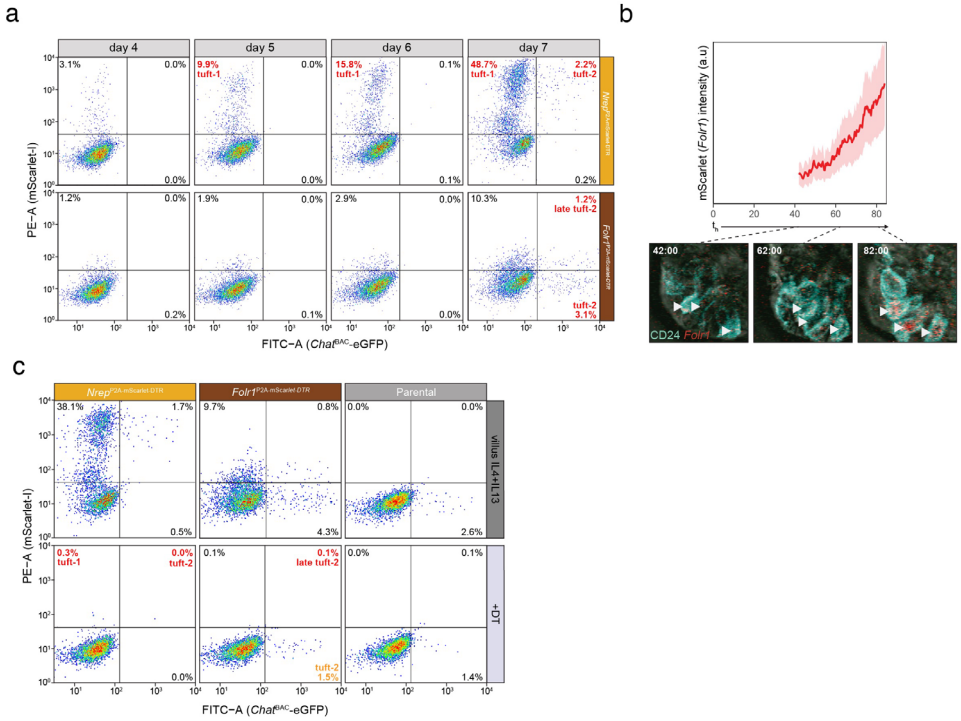
Supplementary Figure 5.6: Type 2 cytokine treatment hardly yields tuft-2 cells *in vitro*. Related to Figure 5.4. **a)** Schematic representation of experimental setup for *in vitro* tuft cell induction. *Chat*^{BAC}-eGFP;*Nrep*$^{P2A\text{-}mScarlet\text{-}I}$ organoids were treated by various regimens (#1- #4) and evaluated by flow cytometry on the 7th day after seeding. **b)** Percentage of mScarlet-I+ and GFP+ cells per inidcated treatment (panel a). Data are represented as mean ± SD (ANOVA mScarlet+ cells $p < 0.01$, ANOVA GFP+ cells not significant, Tukey HSD test). **c)** Gene set enrichment analysis of tuft-1 and tuft-2 signatures[17] in mice infected with *Nippostrongylus Brasiliensis*. Bulk RNA-sequencing data of infected and control mice was obtained from E-MTAB-9183[207]. **d)** Single-cell RNA sequencing profiles of tuft cells isolated from mice infected with *Heligmosomoides polygyrus* (obtained from GSE92332[17]) highlighted in UMAP of tuft cell cluster from integrated mouse *in vivo* single-cell transcriptome dataset (Figure 5.1d). **e)** Proportion of post-mitotic tuft-1 and tuft-2 subtypes during infection with *Heligmosomoides polygyrus* (obtained from GSE92332[17]). Data are represented as mean ± SD. Data points represent individual mice.



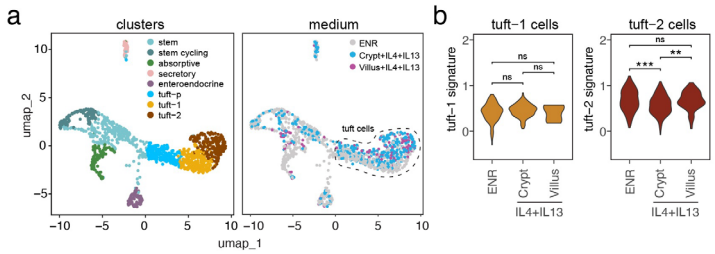
Supplementary Figure 5.7: Live imaging of *Chat*^{BAC}-eGFP;*Nrep*$^{P2A\text{-}mScarlet\text{-}I}$ organoids with anti-CD24. Related to Figure 5.4. **a)** Violin plot showing Cd24a (CD24) expression in organoid scRNA-seq dataset. **b)** Boxplot depicting timing of first detectable *Nrep*$^{P2A\text{-}mScarlet}$ fluorescence signal relative to first detectable anti-CD24 signal. **c)** Representative timelapse sequence showing brightfield images overlayed with anti-CD24 (blue) and *Nrep*$^{P2A\text{-}mScarlet}$ (red) fluorescence signals. Here, CD24 signal was detectable before division, while *Nrep*$^{P2A\text{-}mScarlet}$ was first detected only after division. **d)** *Nrep*$^{P2A\text{-}mScarlet\text{-}I}$ fluorescence signal intensity over time in rare dividing tuft progenitor cells (n = 6 cells). Tracks are synchronized such that t = 0 marks division timepoint (dashed line). **e)** Frequency of cells dividing while positive for CD24 or for *Nrep* during live microscopy experiments (n = 72 cells).



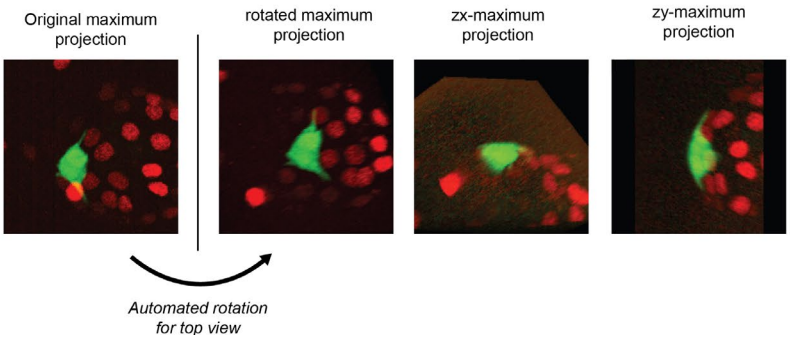
Supplementary Figure 5.8: Tuft cell maturation to tuft-2 phenotypes depends on BMP signaling. Related to Figure 5.5. **a)** Fold-change induction of *Chat*^{BAC}-eGFP+ cells after addition of single villus/autocrine-inspired factors compared to cytokine-pretreated organoids without ENR (condition #2 in Figure 5.5ab). Data are represented as mean ± SD (ANOVA $p < 0.0001$, Tukey HSD test). **b)** Expression level of BMP ligands, receptors and BMP responsive genes in indicated cell types extracted from the integrated *in vivo* mouse small intestine single-cell transcriptome dataset (Figure 5.1d). Average expression is represented by dot color. Percentage of expressing cells is denoted by dot size. **c)** Gene set enrichment analysis of tuft-1 and tuft-2 signatures[17] in bulk RNAseq dataset of intestinal epithelium depleted for *Bmpr1a* (*Vil-CreERT:Bmpr1a*$^{FL/FL}$) vs. control mice[210]. For current analysis, RNA data from multiple diets were included in the comparison between *Bmpr1a* knockout and control groups. Upon perturbed BMP signaling, tuft-2 signature genes are the most downregulated.



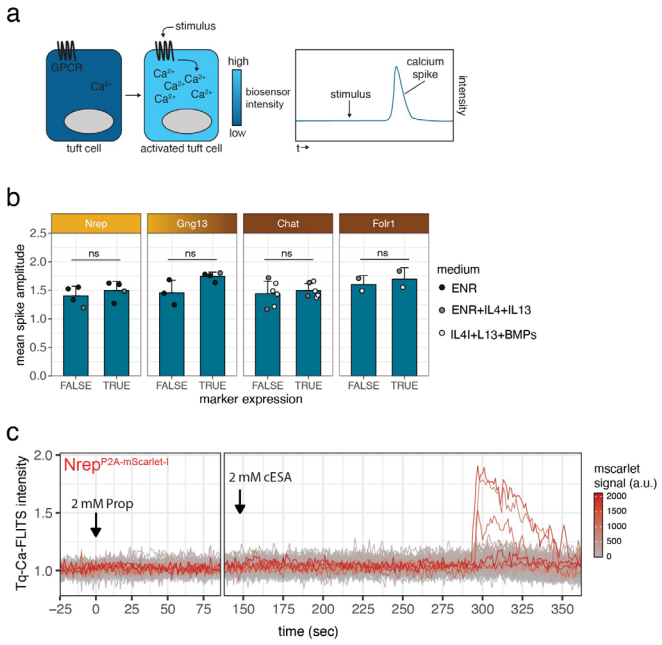
Supplementary Figure 5.9: Temporal dynamics of tuft cell subtype-specific markers in villus medium and depletion with DT. Related to Figure 5.5. **a)** Representative flow cytometry time course analysis of fluorescent cell populations in *Chat*^{BAC}-eGFP;*Nrep*$^{P2A\text{-}mScarlet\text{-}I}$ and *Chat*^{BAC}-eGFP;*Folr1*$^{P2A\text{-}mScarlet\text{-}I}$ organoids treated with villus-inspired maturation regimen (Figure 5.5ab #3). Combined data of 3 biological replicates are shown, 5000 cells are shown per plot. **b)** Top: mScarlet(*Folr1*) signal over time as determined by live imaging in organoids treated with the villus-inspired regimen indicated in Figure 5.5a. Bottom: microscopy stills from indicated timepoints (red: mScarlet(*Folr1*); cyan: anti-CD24). Imaging was started after addition of cytokines on the 3rd day after seeding (n = 10 cells from a representative experiment). White arrows indicate cells that are mScarlet(*Folr1*)+ at the end of timelapse microscopy. **c)** Representative flow cytometry analysis of fluorescent cell populations in *Chat*^{BAC}-eGFP;*Nrep*$^{P2A\text{-}mScarlet\text{-}I\text{-}DTR}$, *Chat*^{BAC}-eGFP;*Folr1*$^{P2A\text{-}mScarlet\text{-}I\text{-}DTR}$ and parental no-knockin control organoids (*Chat*^{BAC}-eGFP) treated with villus-inspired maturation regimen (Figure 5.5ab #3) in the presence or absence of Diphtheria toxin (DT). Combined data of 3 biological replicates are shown, 3000 cells are shown per plot.



Supplementary Figure 5.10: scRNA-seq of organoids in villus-inspired medium. Related to Figure 5.5. **a)** Transcriptomic analysis of single cells in organoids upon induction with cytokines in ENR or villus-inspired culture conditions. Left, UMAP of organoid cells enriched for tuft cells by FACS. Cell types are annotated by color. Right, same UMAP but cells are colored by medium condition. **b)** Violin plots depicting distribution of relative levels of tuft signature scores in tuft-1 cells (left) or tuft-2 cells (right) originating from different culture conditions (ns not significant, ** p-value < 0.01, *** p-value < 0.001, Wilcoxon rank sum test).



Supplementary Figure 5.11: Dynamic tuft cell protrusion analysis. Related to Figure 5.6. Stills exemplifying automated rotation algorithm used to visualize the position of tuft cell protrusions within the organoid epithelial layer. Protrusions wrap around the basal side of the organoid. Green: *Chat*^{BAC}-eGFP, Red: H2B-mScarlet-I.



Supplementary Figure 5.12: Calcium signaling dynamics in organoid monolayers following stimulation with cESA and propionate. Related to Figure 6. **a)** Schematic of tuft cell activation measurements with Tq-Ca-FLITS calcium biosensor. **b)** Bar diagram depicting mean amplitude of calcium spike following stimulation with 1.5 mM cESA. Per bar, each point represents a separate experiment and is colored per medium used. Data are represented as mean ± SD (not significant, t-test). **c)** Single-cell Tq-Ca-FLITS intensity traces following sequential stimulation with 2 mM propionate, followed by 2 mM cESA. Traces are colored by mScarlet-I signal measured in the corresponding cell.

6 Mechanical cell type patterning of the intestinal crypt

Secretory cells in the intestinal epithelium lie interspersed between non-secretory cells, most prominently in the stem cell zone where the niche-maintaining Paneth cells are each surrounded by stem cells. This pattern is generally assumed to be caused by Delta-Notch lateral inhibition in which secretory cells inhibit secretory fate in their neighbors. Lateral inhibition is responsible for similar patterns throughout animal development, but the dynamic, self-renewing nature of the intestinal epithelium poses major challenges to the stability of such patterns. Mathematical modelling of lateral inhibition in the presence of cell rearrangements indicates that stable interspersed patterning is only possible in almost static tissues. Experiments and cell tracking in intestinal instead reveal that the intestinal epithelium is highly dynamic, with cell divisions inducing extensive rearrangements of nearby cells at measured rates that are inconsistent with lateral inhibition as the main driver of the maintenance of secretory cell patterning. Rather, interspersed patterning arises through spatial rearrangements of adjacent secretory cell pairs. Secretory cells brought together by photo-ablation of intervening stem cells show that cell mechanics alone can rapidly reestablish interspersed patterning through the active shrinking of secretory-secretory interfaces. Hence, differences in mechanical properties between cell types, rather than lateral inhibition of cell type through signaling, induce the 'checkerboard'-like patterns in a highly dynamic epithelium like the intestine.[1]

[1] This chapter is based on a manuscript in preparation by M.A. Betjes, L. Dubbink, S.J. Tans, J.S. van Zon.

6.1 Introduction

Secretory cells in the intestinal epithelium are positioned isolated from each other in between non-secretory cells (**Figure 6.1A)**. In the stem cell zone at the crypt bottom, Paneth cells are interspersed between stem cells in a so-called 'checkerboard-pattern'[24,169,237]. Above this zone, secretory (progenitor) types, like goblet, enteroendocrine and tuft cells, are scattered between absorptive cells[45,195,238]. This arrangement of secretory cells can modulate cell function and intercellular signaling in crucial ways. Paneth cells, for instance, produce stemness maintaining niche factors, whose function often depends on direct contact between Paneth and stem cells, either because these factors are presented on the surface [169,239] (Notch ligands) or because their insolubility renders them membrane bound[240] (Wnt proteins). Indeed, the size of contact area between Paneth and stem cells directly affects maintenance of stemness and reduced contact area leads to a marked decrease in regenerative capabilities[241].

Such patterns where cells of one type are separated from each other by cells of another type, are ubiquitous in epithelia across different organs and organisms. Many of these patterns, from the vertebrate inner ear to the insect eye, are generated through notch signaling. Here, a cell type that expresses notch ligands inhibits differentiation towards that cell type in its direct neighbors, which express the notch receptors[242].

Similarly to these systems, in the intestinal crypt secretory cells express notch ligands, while the surrounding stem cells and absorptive progenitors, express notch receptors[169] (**Figure 6.1A**). Indeed, the notch signaling axis is key in secretory versus non-secretory cell fate specification and stem cell maintenance[12,157]. Although an explanation for how exactly the Paneth-stem pattern is achieved has of now remained out of reach[162], it is thus often assumed or speculated that the Paneth-stem pattern is established through a lateral inhibition mechanism[16,243], where Paneth cells inhibit differentiation in their stem neighbors. The scattered distribution of the other secretory cell types in the intestine, such as enteroendocrine and goblet cells, has similarly been attributed to lateral inhibition[16,244,245].

In contrast to the canonical examples of lateral inhibition in sensory epithelia[242], the small intestine epithelium is rapidly renewing. Stem cells and absorptive progenitor cells therefore have to divide constantly and are rapidly pushed out of the crypt[86], which could potentially force cell rearrangements[246,247] and disturb the checkerboard pattern (**Figure 6.1A**). Given that secretory cell lifetimes are long compared to the average cell cycle duration[88], it remains an open question if a lateral inhibition mechanism can overcome these disruptions and is able pattern such dynamic epithelia[248,249].

Theoretical models of lateral inhibition in the intestinal stem cell niche indicate that it can give rise to checkerboard patterns, but these models pay little[243] or no[157,250,251] attention to the possibility of cell division or rearrangements. In 3D models of the intestinal crypt[252-255] or intestinal organoids[256,257] that do include cell division and movement, it is often noted that

the implementation of lateral inhibition should give rise to a checkerboard pattern, but the resulting patterning is not quantified. Generally, lateral inhibition does not seem to produce a strong alternating pattern in these models.

Further complications to a lateral inhibition model arise from the fact that, at least in organoids, all secretory cells (including Paneth cells) seem to undergo at least one division after fate commitment[88,258]. Although many of the sister pairs will separate from each other during the division process, at least half of them seem to remain neighbors[45,259], thereby disturbing the patterning. Recent work from our lab suggests that the remaining secretory sister neighbor pairs do tend to move apart on a longer time scale[88]. Similarly, for goblet progenitors in the colon epithelium, that have long been known to divide[260], simulations have suggested that they 'actively disperse' in some unknown to break up goblet clones[261].

Experimentally, the robustness of the intestinal stem cell niche during photoablation experiments[262] or after the conditional ablation of Paneth[263] or stem cells[264] suggests a role for cell movement in pattern maintenance, but little work has been done on the mechanical basis of these phenomena. Studies of intestinal cell mechanics have thus far mostly focused on the geometry and forces governing the formation of the crypt and villus domains[219,265,266], but not on how cell types are patterned within these domains.

Here, we use long-term live imaging of intestinal organoids combined with cell tracking to study the patterning of secretory cells in real time. We find that the intestinal crypt is indeed a highly dynamic environment with cell divisions driving fast rearrangements. Moreover, a simple theoretical model shows that it is impossible for lateral inhibition alone to maintain a pattern in these conditions. Instead, our cell tracking indicates that the fast separation of secretory neighbors is key in pattern establishment. Using photoablation experiments, we show that the different interfaces between the cell types have distinct mechanical properties that actively drive T1-transitions to separate secretory cells. Moreover, Paneth cells are stiffer and less mobile than stem cells in response to nearby cell extrusions, thereby conferring extra stability to the pattern. Our study establishes a new mechanical paradigm of checkerboard-pattern formation that can overcome the challenges posed by the constantly self-renewing nature of the intestinal epithelium. Considering epithelia more broadly, we speculate that cell mechanics and notch signaling act synergistically, thereby conferring robustness across many different contexts, from systems without cell divisions and limited rearrangements to rapidly self-renewing tissues.

6.2 Results

6.2.1 Theoretical model for cellular checkerboard patterning

To quantify the effect of cell rearrangement on the patterning ability of lateral inhibition, we first constructed a simple theoretical model. Our aim here is to model the fraction of secretory neighbor pairs (out of all neighbor pairs), while ignoring higher level spatial structure (e.g. if

the secretory neighbor pairs are part of larger patches). This simplification works when secretory cells are well-mixed throughout the epithelium, either randomly or in a checkerboard pattern, which is precisely the space of configurations we are interested in. By ignoring higher order structure, we can model the creation and dissolution of secretory cell pairs simply on the basis of the overall fraction of secretory cells and secretory neighbors in the system. We can thus avoid implementing a full model of the epithelium that explicitly describes the movement of all the cells (**Appendix**).

The two main processes affecting secretory patterning are cell rearrangements and the differentiation of secretory cells **(Figure 6.1B)**. Cell rearrangements can create and force apart secretory cells through T1-transitions, where two neighbors move apart to allow a new neighbor pair to form[267], thereby randomizing the pattern. Differentiation can create order, because lateral inhibition generally ensures that a new secretory cell will not emerge next to another secretory cell. Secretory differentiation time is key in determining lateral inhibition efficiency: even if the secretory progenitor is perfectly patterned by lateral inhibition, the mature secretory cell that emerges from it can already, due to rearrangements, have assumed a new position incompatible with the desired checkboard pattern. Similarly, any divisions during the progenitor stage also disturb the patterning. Beyond these two key processes, our model also allows the incorporation of cell extrusions and divisions of both secretory and non-secretory cells (**Supplementary figure 6.1)**.

We characterize the resulting degree of patterning in our model (φ) by the fraction of secretory neighbor pairs over the expected fraction of such pairs when cells would be arranged randomly. Here, $\varphi=0$ if the secretory cells are perfectly separated in a checkerboard pattern, $\varphi=1$ if cells are randomly distributed and $\varphi>1$ if the secretory cells are clustered **(Figure 6.1C)**.

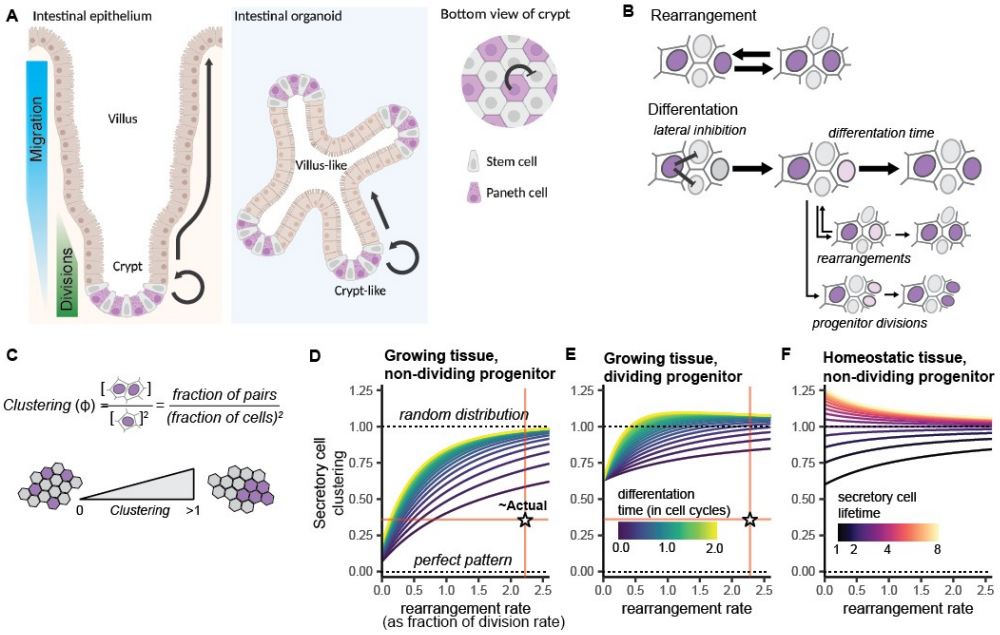


Figure 6.1: Lateral inhibition-mediated checkerboard patterning in tissues with cell rearrangements. A) Schematic of the intestinal epithelium and intestinal organoids. The epithelium consists of crypt and villus(-like) domains. Cell divisions occur in the crypt, replenishing cells lost in the villus. At the bottom of the crypts, Paneth cells (purple) form a checkerboard pattern with Paneth cells interspersed between stem cells. Paneth cells inhibit differentiation of surrounding stem cells. **B)** Schematic explanation of a minimal mathematical model of secretory cell clustering: rearrangement randomizes the established interspersed pattern by bringing together and separating secretory cell pairs. Differentiation with lateral inhibition tends to create isolated secretory cells improving patterning, but rearrangements and progenitor divisions during the differentiation process counteract this. **C)** The degree of clustering φ is measured by dividing the fraction of secretory pairs out of all neighboring pairs by the expected fraction if the cells were randomly scattered, which is the square of the secretory cell concentration. When φ=0 secretory cell patterning is perfectly interspersed, φ=1 corresponds to a random distribution, while φ>1 corresponds to clustering of secretory cells. **D)** Patterning in a growing tissue as a function of the rearrangement rates and secretory cell differentiation times (indicated by color). Rearrangement rate and differentiation time are both expressed in terms of the cell division rate, with the inverse of the cell division rate corresponding to the cell cycle time. Proper patterning (φ≈0) can only be achieved in almost static tissues. The star and red lines indicate approximate values of clustering and rearrangement rates based on organoid data. **E)** Same as **D),** but with secretory progenitors undergoing a single division at the start of differentiation. **F)** Patterning in a homeostatic tissue as a function of the rearrangement rates and secretory cell lifetimes (indicated by color). Weakly interspersed patterning (φ <1) is only achieved when secretory cells are short-lived.

6.2.2 Lateral inhibition is insufficient for patterning at moderate rearrangement rates

We first apply our model to the case of a growing tissue. This resembles conditions in the organoid crypt, which generally expands during culture, or the in-vivo transit-amplifying zone located above the stem cell zone. In this case (as in the others presented below), secretory neighbor clustering will reach a steady-state level. We probe how this steady state depends on our key parameters; the differentiation time that governs lateral inhibition

efficiency and the rearrangement rate. We define all these parameters relative to the division rate in the system.

We find that not only increasing rearrangement rates but also prolonged differentiation times strongly perturb pattern formation (**Figure 6.1D**), with the latter creating a close to random patterning ($\varphi > 0.75$) even at modest rearrangement rates, i.e. at the same rate as cell divisions. But even for very short differentiation times (less than half a cell-cycle, corresponding to the time required for onset of expression of early markers of secretory fate[157]), almost no cell rearrangements are allowed for a pattern to emerge.

In our second scenario, we let the secretory progenitor undergo a final cell division after commitment. This reflects recent evidence for the existence of dividing secretory progenitors, at least in organoids[88]. This progenitor division counteracts the lateral inhibition by creating a new secretory pair during differentiation. In this case, close to random distributions of secretory cells (phi>0.75) emerge even in almost immobile tissues (**Figure 6.1E**). Assuming that the daughter cells tend to separate during the division process, as previously described in intestinal crypts and organoids[45,77,88], does little to improve patterning, as this sister separation also produces new random neighbor interactions **(Supplementary figure 6.2A)**.

Lastly, we consider a non-growing homeostatic system like the in vivo stem cell zone at the crypt bottom, where cell divisions must be compensated by cells leaving the system. Following established models for stem cell dynamics we treat these exiting cells like extrusion events[142,268], with their absence creating new neighboring pairs within the remaining population. In this scenario, we do not let secretory cells divide to allow for the best possible patterning. We find that here that in addition to the differentiation time **(Supplementary figure 6.2B)**, the rate of secretory cell extrusion becomes even more important for cell patterning. Even under the most favorable conditions, like almost instantaneous differentiation (4 hours) from stem to secretory types, we find that checkerboard patterns can only be produced by assuming unrealistically short secretory lifetimes of one cell cycle or shorter **(Figure 6.1F)**, while in vivo evidence suggests that Paneth cells can live for weeks[269]. This can be explained by the fact that in homeostasis, as opposed to expanding systems, new secretory cells can only emerge at the rate at which existing secretory cells are lost. The secretory cell lifetime thus directly limits the influence of differentiation through lateral inhibition on the pattern. Even in a model where stem cells manage to avoid creating secretory neighbors while leaving the system, a strong checkerboard patterning only emerges when cells almost never rearrange **(Supplementary figure 6.2C)**.

6.2.3 Organoid crypts exhibit frequent division-driven cell rearrangements

Our minimal model indicates that achieving a checkerboard pattern through lateral inhibition is either impossible or only possible at very low rearrangement rates, raising the question of how readily cells rearrange in the intestinal epithelium. To address this, we examined cell

rearrangement through live imaging of intestinal organoids, using fluorescently labeled nuclei to automatically track >100 cells for over 40 hours in intestinal organoids (**Methods**). We first focused on cell movements during cell division as the intestine's rapidly dividing nature is one of its key distinguishing features compared to other epithelial systems that exhibit 'checkerboard'-type patterns. Moreover, cell division in the intestinal epithelium is also accompanied by strong movement of the cell nucleus, moving first to the apical side before reintegrating basally after cell division[45,259], which could possibly contribute to local cell rearrangements.

We first compared the position of mother cells before division (1 hour before metaphase) to the position of their daughter cells (1 hour after). Daughters often move considerably as a result the division process, with relative displacements close to the median distance between two neighboring cells (~8 um, **Figure 6.2A)**. The typical distance between sister cells after division exceeds the median neighbor distance, indicating that sister cells move substantial distances away from each other during the division process (**Figure 6.2B)**. When we examined the displacements of both daughter cells relative to the position on the mother cell, we observed three different classes of sister cell displacement (**Figure 6.2C)**. First, we find sister pairs that separate by both moving away from the mother's position (18%, sisters move >1.5 cell diameter apart and both more than half that distance away from their mother's position), which was described previously as the dominant mode of sister separation[45,259]. However, we found many cases of sister separation that showed a different displacement pattern. Specifically, we found cases of sister separation where one daughter remained at the mother cell position, while its sister moves away (16%, sisters move >1.5 cell diameter apart, but with one cell remaining close to the mother's position). Finally, we find cases where the sisters remain neighbors and stay close to their mother's position (65%, sisters that move <1.5 cell diameter apart).

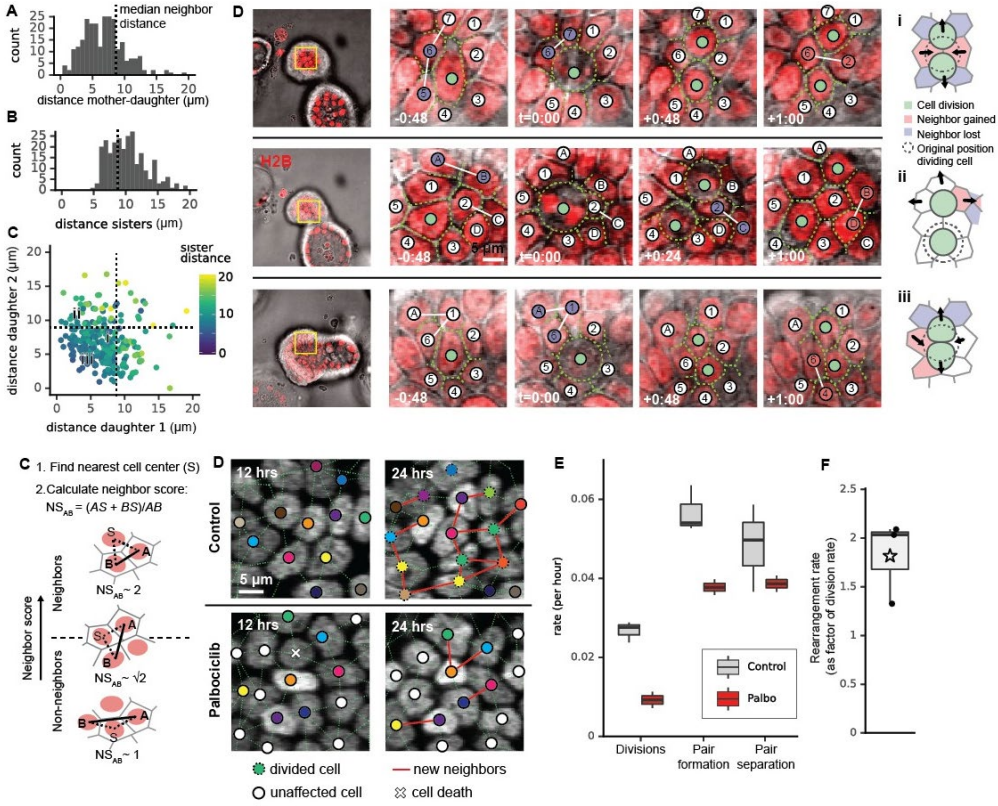


Figure 6.2: Intestinal epithelial cell rearrangements driven by cell division. A) Histogram of the distance between daughter (1 hour after division) and mother cell position (1 hour before division). **B)** Histogram of the distance between sister cells (1 hour after division). **C)** Distance between daughter and mother position, as defined in A), for sister pairs. Color denotes sister distance after division. **D)** Three different modes of cell division-driven cell rearrangement. Dividing cells are marked green, with surrounding cells given a number. Cells located further are labelled with letters. Lost neighbor pairs are marked in blue, new pairs in red. i) Sister cells both move away from the mother's position, creating a new neighbor pair in between them. ii) Sisters separate but only one moves away from the mother cell position. Rearrangements occur around the position of the far-moving sister. iii) The sisters stay close to each other. Rearrangements are still induced by the elongated shape of the sister pair compared to that of mother cell. Cartoons give a schematic overview of the effect of the different modes of cell division on their neighbor. Arrows show cell movement induced by divisions. **E)** Schematic automated detection of neighbors, based on relative positions of cell nuclei. **F)** New neighbor pairs emerging in organoid crypts over 12 hours for control and cell-cycle inhibited (palbociclib) conditions. The dividing tissue has many more rearrangements, even for cells that have not divided themselves. **G)** Rates of cell division, rates of new neighbor pair formation and the rates at which neighbor pairs disappear (move apart), both in control and palbociclib-treated organoid crypts. Rearrangements were only considered when not driven by a division of either of the cells in the neighbor pair. Palbociclib significantly reduces both the division rate and the rearrangement rates for cells. **H)** Rearrangement rate (here taken to be the pair formation rate) in terms of the cell division rate for control organoids. The average (star) can be compared to our model in Figure 6.1D and E.

Next, we examined the impact of these three different sister separation classes on cell rearrangements in adjacent non-dividing cells, using the transmitted light signal to resolve cell outlines (**Figure 6.2d)**. These effects are of special interest to us as mature secretory cells will generally not divide themselves but are still subject to the influence of divisions of nearby non-secretory cells. Divisions where both daughter cells displace away from the mother's position, typically lead to creation of a new neighbor pair from cells that were previously on different sides of the mother cell, while the daughter cells separate neighboring cell pairs in the direction of their movement (**Figure 6.2D,i**). This is consistent with the proposed role of this mode of sister displacement as a main driver of epithelial fluidization[45,246]. Interestingly, however, we find that the sister pairs where one or both sisters remain close to the mother cell's position can still generate substantial rearrangements. When one sister moves away from the mother's position, we do not see a new neighbor pair forming between the sister cells, but instead see rearrangements occur further away caused by the reintegration of the displaced sister (**Figure 2D,ii**). When both sisters remain adjacent and close to the mother's position, this causes rearrangements among the adjacent cells due to the elongated shape of the sister pair along the division axis compared to the mother cell (**Figure 2D, iii**). We note that this latter displacement pattern was observed to not directly cause cell rearrangements in a diverse range of metazoan epithelial sheets[246,270] and speculate that this might arise here due to the highly packed and columnar nature of the intestinal epithelium. Crucially, all three different classes of sister displacements thus have the capacity to force cell rearrangements in nearby cells, showing the importance of cell division in driving intestinal cell rearrangement.

We next sought to experimentally quantify the rearrangement and division rates of the intestinal epithelium to compare with the results of our mathematical model. We automated the detection of rearrangements by identifying neighboring cell pairs based on the centers of their nuclei (**Figure 6.2E, Methods)**, allowing estimation of the rearrangement rates without need for a full segmentation of the cell surface. Visual inspection of neighbor pairs detected in this way shows that the dividing epithelium is highly dynamic: within a 12-hour window virtually all cells acquired new neighbors (**Figure 6.2F)**. This was the case both for cells that have undergone division and non-dividing cells. We then examined the effect of cell divisions on the rearrangement rate. To that end, we also imaged organoids while inhibiting the cell cycle using Palbociclib[235]. Without divisions, the epithelium is much more static with many cells retaining the same neighbors over a 12-hour period. However, rearrangements do occur, for example in the vicinity of cell deaths (**Figure 6.2F)**.

We then quantified the rearrangement rate by measuring the rate at which novel neighbor pairs emerged in the crypt and the rate at which established neighbor pairs separated (**Figure 6.2G,H**). Here, we excluded neighbor pairs that formed or disappeared because one cell, or both, divided, thus only considering spontaneous rearrangements or those driven by cell divisions elsewhere in the epithelium. To examine the impact of cell divisions on the

observed rearrangements, we also quantified the cell division rate, using a neural network trained to identify cell divisions from nuclear fluorescence signals (**Methods**)[235]. Palbociclib treatment indeed decreases the division rate in the crypt from 0.03 to 0.01 divisions per cell per hour (**Figure 6.2G**). This leads to a significant reduction in the rate at which neighbors separate, from 0.06 to 0.04 separations per neighbor pair per hour (**Figure 6.2G)**, with a similar decrease seen for the rate of neighbor pair formation. These results thus demonstrate the role of cell divisions in forcing cell rearrangements. These measurements also allow estimation of the ratio between the rearrangement (here taken to be equal to the rate of pair formation) and division rate during normal homeostasis. We find that the former is significantly higher (~2-fold) than the latter, showing the importance of considering not only cell division and differentiation but also at cell rearrangement when looking at the ability of lateral inhibition to pattern a tissue (**Figure 6.2H)**. In our mathematical model, such high rearrangement rates exclude checkerboard pattern formation even under the most favorable assumptions (**Figure 6.1D,E)**. Overall, the ability of cell divisions to induce rearrangements suggests that lateral inhibition alone might be incapable of maintaining patterning across a wider range self-renewing (and thus constantly dividing) epithelia.

6.2.4 Interspersed patterning arises by rapid secretory neighbor separation.

In order to identify possible mechanisms for the interspersed patterning of secretory cells, we analyzed an existing intestinal organoid dataset, where individual cells were tracked for over 60 hours after which they were stained for multiple cell type markers, providing us with both dynamic spatial and cell type information[88]. We first checked if the in vivo checkerboard patterning is recapitulated in organoid crypts. By looking at the fraction of neighbor pairs of all cell type combinations, we indeed see that Paneth cells are less often adjacent than expected by chance (**Figure 6.3A, B)**, with the same holding also for enteroendocrine cells and mucus-producing cells that represent Paneth and Goblet progenitors[88]. Secretory pairs consisting of two different secretory types were also comparatively rare, indicating that any mechanism for interspersed secretory patterning acts also between different secretory cell types. As expected, Paneth cells and mucus-producing cells were enriched for contacts with stem cells, consistent with their location in the crypt. Stem cell contacts are also enriched for enteroendocrine cells, which were also predominantly located in the crypt, probably reflecting their immature nature in organoid culture[14,15]. Conversely, these secretory cell types showed no contact with enterocytes, indicative of their exclusion from the villus compartment and thus confirming that our approach can recover the known tissue architecture of the intestine **(Figure 6.3B)**. The average secretory cell clustering score within the crypt region was φ=0.33 (**Figure 6.3C)**, consistent with secretory cell separation. Similar scores were found when looking at the patterning of Paneth, enteroendocrine or mucus-producing cells specifically (**Figure 6.3C)**. This observed quality of interspersed secretory cell patterning is much higher than what could be achieved with lateral inhibition in our theoretical model at the measured rearrangement rates (φ>0.75, **Figure 6.1E)**.

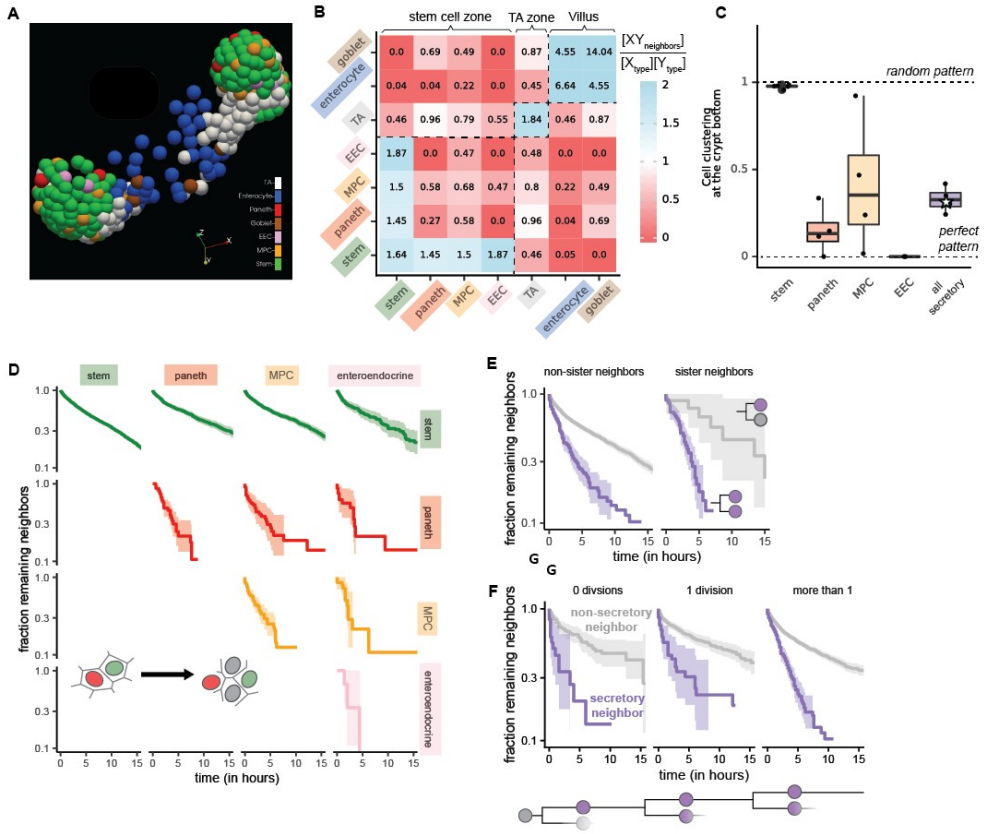


Figure 6.3: Rapid separation of secretory cell neighbors. A) Reconstructed intestinal organoid with cell positions (spheres) determined by cell tracking and color representing cell type as determined by antibody staining, with color key defined in B). **B)** The clustering of different cell type pairs in the organoids (n=5), with blue indicating cell type pairs more often adjacent than expected and red instead separated. The dotted squares indicate the stem, transit-amplifying (TA) and villus compartment that are also visible in A). Within the stem cell compartment, Paneth cells, mucus producing cells (MPCs) and enteroendocrine cells (EECs) are found separated from one another. **C)** The cell clustering within the stem cell compartment for stem cells, the different secretory types and all secretory cells pooled together. The dotted line denotes a random rearrangement, zero clustering denotes perfectly separated secretory cells. The average (star) is compared against the model in Fig. 1D,E. **D)** Fraction of cells still in contact, as function of time after becoming neighbors, for different cell type neighbor pairings in the stem cell zone. The fraction of stem and secretory cells remaining neighbors after 10 hours is considerable (~30%), while all pairings of secretory cells separate much faster. **E)** Secretory neighbors (purple line) separate faster than secretory-stem pairs (grey line), both when they are sisters or not. **F)** As in E, but split according to the differentiation stage of the secretory cell, as quantified by the number of divisions since cell fate commitment. Secretory neighbor separation is present both for differentiating and fully differentiated cells.

After confirming that the organoid recapitulates the in vivo checkerboard pattern by having comparatively few secretory neighbor pairs, we wanted to see what happened to those pairs that did form. We analyzed the duration of neighbor pairings, starting from the moment both cells came into contact, using survival analysis, which allowed us to rigorously deal with

observations that were cut short because one of the cells in the pair divided, died or was lost to tracking[235]. By comparing secretory-secretory pairs to stem-secretory pairs in the crypt, for all secretory types, we found that secretory neighbor pairs remained together for significantly shorter duration than neighbor pairs with one stem cell (**Figure 6.3D)**. While Paneth-stem neighbor pairs remained together for ~12 hours, we saw that most Paneth-Paneth pairs separated within 4 hours, with similar short duration of neighbor contacts seen for all secretory type combinations in the crypt.

Secretory neighbor can be formed by secretory cell differentiation, by existing secretory cells coming into contact through rearrangements, or by two secretory sisters emerging next through one another following division of their mother. We found that adjacent secretory sister cells separate on the same rapid timescale (4 hours) as secretory neighbor pairs that were not sisters (4.5 hours, **Figure 6.3E**). Moreover, secretory sisters separated more rapidly than rare sister pairs consisting of one secretory and one non-secretory cell (~15 hours), excluding that adjacent sister cells separate more readily than neighboring cells that are not sisters. Overall, these observations indicate that separation of secretory cells occurs does not rely on their degree of relatedness or their recent emergence through division.

We next sought to see how early in the differentiation process we observe the tendency of secretory cells to separate. Using the lineage information and backward propagation of cell type as determined by antibody staining[88], we can identify the cell cycle in which commitment to secretory type takes place, as the last cell cycle in which the cell had a non-secretory sister (**Figure 6.3F**). Given the short time since inferred induction of secretory type and the fact that these cells divide at least once to generate offspring of the same secretory type, the cells at this first cell cycle likely reflect a secretory progenitor state, rather than a terminally differentiated cell. However, even this shortly after commitment, we find that secretory cells separate more rapidly from secretory neighbors than from non-secretory neighbors (**Figure 6.2F**). While this observation might point to separation even of newly differentiated secretory cells, it could instead also reflect secretory differentiation through lateral inhibition, with cells assuming secretory fate exactly because they move away from adjacent secretory cells, rather than the reverse, something which we cannot infer from our data. However, in the subsequent cell cycle, when we are certain the secretory progenitor has committed, we find similar rapid separation of secretory cells (5 hours average duration) as for mature secretory cells (4.5 hours, defined as cells that have undergone more than two divisions since commitment or have not divided in our ~60 hours experimental window). Overall, these results indicate that the ability of secretory cells to rapidly separate from one another does not emerge gradually as the cell terminally differentiates but instead commences shortly after or possibly even during secretory commitment.

We then examined the impact of rapid secretory cell separation on our minimal mathematical model. When we increased the rate of secretory cell separation 3-fold compared to separation of other cell type pairs, we could readily achieve the observed interspersed secretory cell

patterning (φ=0.3, **Supplementary figure 6.3)**. Interestingly, in the presence of fast secretory neighbor separation, the dependence on differentiation time largely disappears from our model as lateral inhibition efficiency ceases to be of concern: the pattern is now driven by cell movement not by cell signaling. We wanted to see if this predicted independence of pattern quality on differentiation time holds in our experimental data as well. In our model differentiation time is defined as the time between fate commitment and the marker used for quantifying the pattern. At early timepoints in the timelapse data the lineage information allows us to identify very early progenitors that will later give rise to mature secretory cells, as if we had access to a very early marker. Quantifying the pattering at these early timepoints thus gives an estimate of the patterning at very short differentiation times. Inclusion of early progenitors in our quantification would improve a lateral inhibition driven pattern as early secretory progenitors are in that case patterned better than mature cells because rearrangements would have had less time to do their randomizing work (**Figure 6.2D, E)**. In contrast, we see no difference in pattern quality when including early progenitors, as expected from patterning driven by secretory cell separation (**Supplementary figure 6.4)**. in conclusion, these observations indicate that rapid separation of secretory cell neighbors is the main driver of spatial secretory cell type patterning.

6.2.5 Secretory cell separation driven by shrinking of cell interfaces

To further probe the mechanism of secretory cell separation, we then sought to externally perturb cell type patterning in the crypt to bring secretory cells together and examine how such situations are resolved. Specifically, we used laser ablation by a 800nm two-photon laser (**Methods**) to induce cell extrusion[271] of stem cells in organoid crypts that were positioned between two secretory cells, resulting in the formation of a secretory neighbor pair (**Figure 6.4A, B).** We identified secretory cells using a conjugated antibody live-cell marker for CD24, which marks all secretory cell types in the crypt[263], with the majority of these cells (~70%) identified as mature Paneth cells by the presence of granules visible by transmitted light. We use an organoid line with fluorescently labeled E-cadherin (E-cadherin-mNeonGreen)[271], which labels the cell membrane, enabling us to directly follow how cell shapes and interfaces change in time in response to this perturbation.

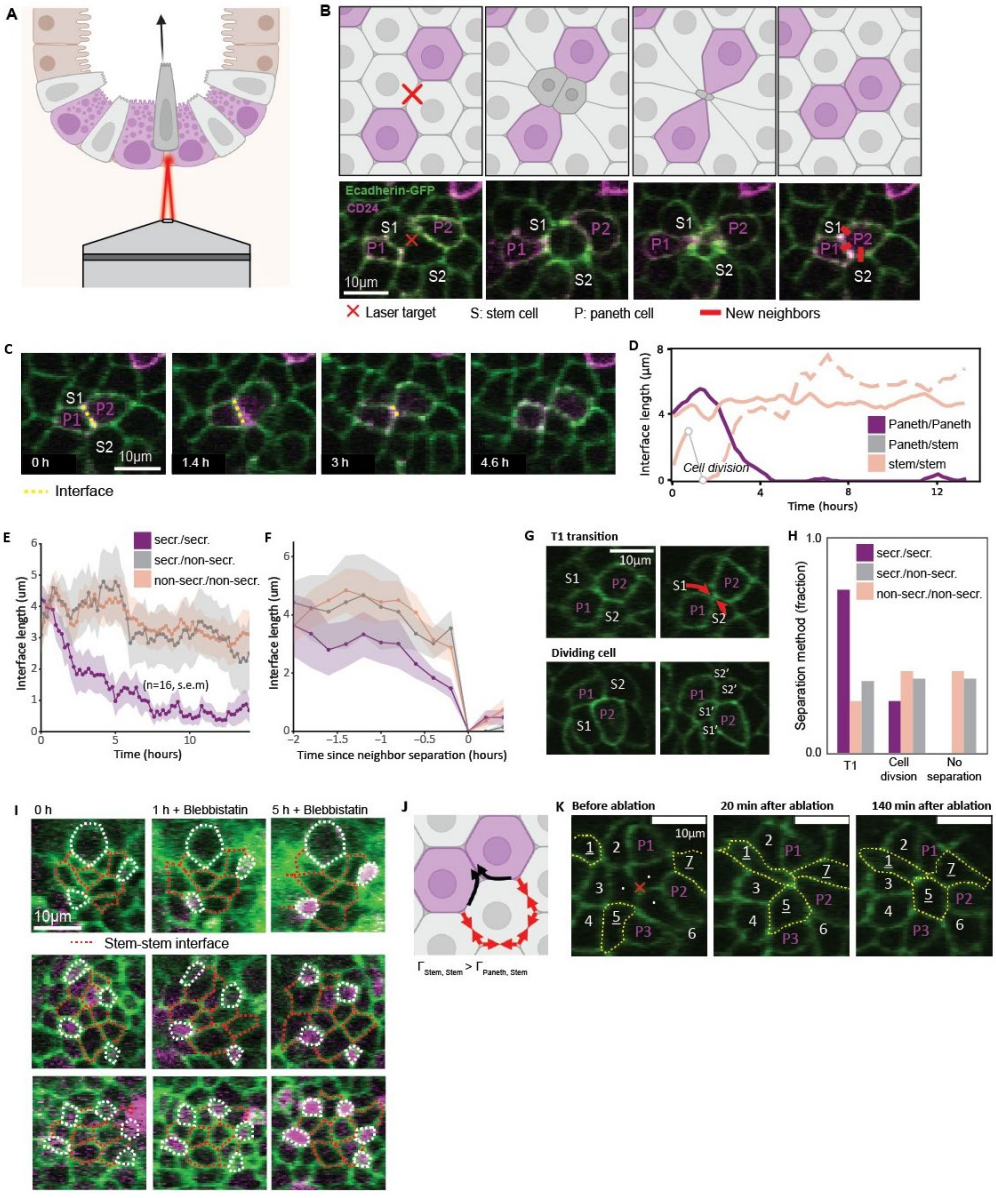


Figure 6.4: Different cell interface dynamics drive separation of secretory cells. A) Schematic of the photoablation procedure. Stem cells are targeted basally by a two-photon laser upon which they extrude from the epithelium apically. **B)** Laser-ablation induced cell extrusion, from left to right: two stem cells are targeted for extrusion by centering the laser pulse basally at their interface; a contractile ring forms; the two cells are extruded; new Paneth-Paneth and Paneth-stem interfaces dare formed. **C)** An example of a Paneth-Paneth interface shrinking over a ~4 hour time course. **D)** Interface length over time for the Paneth-Paneth interface in C) and the two newly created Paneth-stem interfaces. Only the Paneth-Paneth interface shrinks. **E)** Average interface lengths for n=16 photoablation experiments, shaded region is the standard error of measurement. The Paneth-Paneth interfaces shrink strongly over a 7 hour period, while the other interface types remain stable. **F)** Length of different cell interfaces,

for interfaces that disappear during rearrangements following induced cell extrusion, with time relative to the moment of disappearance. The stem-stem and Paneth-stem interfaces seem to separate in a single frame (12 minutes), while Paneth-Paneth interfaces show a gradual decrease. **G)**examples showing that cells can separate by a spontaneous T1-transition or induced by nearby cell divisions. **H)** Paneth-Paneth interfaces (n=16) tend to undergo T1-transitions over the course of the photoablation experiment more often than other interface types. This leads them to undergo separation in all cases, as opposed to stem-stem and Paneth-stem interfaces, which remain present for the full imaging window (15 hours) in around 30% of cases. **I)** 2D-cultured intestinal organoids subjected to blebbistatin treatment which increases the length of stem-stem interfaces (red) and not the of stem-secretory interfaces (white). Apical side is shown. **J)** Schematic model showing how increased contractility can cause secretory pair separation: due to the high contractility of stem-stem interfaces, these shrink in favor of growing stem-secretory interfaces. When two secretory cells share an interface, increasing secretory-stem interface length leads to shrinkage of secretory-secretory interface and their ultimate separation. **K)** Ablation of three stem cells in between 3 Paneth cells, labelled P, does not lead to any Paneth-Paneth pair, because of the strong elongation and migration of nearby stem cells, labelled 1, 5 and 7

In close to half of the cases, stem cell ablation resulted in formation of a secretory neighbor pair. The resulting cell rearrangements also induced new interfaces between these secretory cells and surrounding stem cells, and between these stem cells mutually, which function as a control group to the secretory pairs (**Figure 6.4B**). We characterized the cell separation process by measuring the interface length at the basal side of the organoid between all newly formed neighbor pairs. After extrusion, all newly generated pairs (secretory-secretory, secretory-stem, stem-stem) have similar interface lengths of ~4 μm. The secretory neighbor interfaces then shrink over a 4-hour period until they disappear, then the two secretory cells become separated by a newly formed interface between two surrounding stem cells (**Figure 6.4C-E)**. In comparison, other stem-stem and secretory-stem interfaces show little shrinkage over this period. This observed timescale of secretory interface shrinkage is in good agreement with our previous nuclei-based analysis of secretory neighbor separation under unperturbed conditions. Consequently, the observed cell interface dynamics enable the interspersed secretory cell pattern to reestablish itself mechanically, without any changes in cell type, as we confirmed by an absence of changes in CD24 level.

When then examined if cell interface dynamics during the separation of secretory neighbors differed from those seen during the separation of other cell type neighbor pairs. We therefore compared the dynamics of secretory-secretory, stem-secretory and stem-stem cell interfaces for all interfaces that ultimately disappeared following stem cell ablation, aligned relative to the time of disappearance (**Figure 6.4F**). While secretory-secretory interfaces gradually reduce their length until they lose contact at ~1 μm interface length, consistent with our observations above, stem-stem and secretory-stem average interface lengths remain at ~4 μm to the last frame (12 minutes) before their separation. This suggested that secretory-secretory neighbors separated through different cell interface dynamics than other cell type pairs. When we examined by eye all separating neighbor pairs newly formed by the stem cell ablation, we could indeed distinguish two different separation mechanisms: a stereotypical T1-transition preceded by gradual shrinking of a single cell interface length; and an almost immediate separation of cell pairs invariably accompanied by division(s) of nearby cells

(**Figure 6.4G, Supplementary figure 6.5)**. Separations by T1-transitions were much more prevalent for secretory neighbor pairs (75%) than for other neighbor types (33%), which accounts for the difference in average interface dynamics seen between them. The rate of division-driven separation was similar for all cell type pairs. Moreover, we find that a substantial fraction of newly formed secretory-stem and stem-stem neighbor pairs (33% and 31%, respectively) did not separate, while secretory neighbors always separated (**Figure 6.4G)**, consistent with the notion of a distinct mechanism of separation of secretory neighbors. Our analysis specifically suggests that mechanical differences at secretory-secretory cell interfaces drive the preferential separation of these neighbor pairs, while division-induced cell rearrangement randomly affects all cell pairs.

A potential explanation for the different cell interface dynamics is that interfaces have differential contractility that depends on the cell types connected by the interface. As myosin drives apical cell constriction in the crypt to generate its curved geometry[219,241,265,266], we reasoned that apical myosin-driven contraction might drive differential interface contractility. To test this hypothesis, we inhibited myosin contractility using blebbistatin[241] and analyzed the impact on cell interfaces, using 2D organoid culture to facilitate more reliable measurements of apical cell interface lengths compared to the curved geometry of 3D cultures (**Figure 6.4I)**. Upon Blebbistatin treatment, we saw a strong increase in the apical length of stem-stem interfaces, while secretory-stem interfaces remained largely stable. This indicates that it is the stem-stem interface that is highly contractile, while the stem-Paneth interface is not. This result indicates a mechanism where secretory cells are separated by high stem-stem interface contractility, as this specifically shrinks stem-stem interfaces, at the expense of expanding the stem-secretory interfaces and dissolution of secretory-secretory interfaces (**figure 6.4J)**. A potential benefit of this mechanism over one where high secretory-secretory interface contractility moves secretory pairs apart is that it is immediately active when a secretory-secretory interface is formed and hence does not require a waiting period in which the contractility has to be build up on the new secretory-secretory interface. The immediate initiation of secretory interface shrinkage we see after secretory neighbor formation is thus consistent with this mechanism (**figure 6.4D)**.

Finally, we note that stem cell ablation often did not induce formation of secretory neighbors (~53% cases, n=48) even when ablating as many as three stem cells in between three secretory cells (**Figure 6.4K)**. In those cases, instead of bringing secretory cells together, stem cells deform and elongate to fill in the extrusion site, whereas secretory cells stay rounded and close to their original positions (**Figure 6.4K)**. These observations indicate that secretory cells are thus less prone to moving or deforming than stem cells under the large perturbations of epithelial structure following cell extrusion, which might provide an additional cell mechanical interaction which protects the checkerboard pattern against disruption.

6.3 Conclusions

In this work, we resolve a fundamental question about the intestinal epithelium: can lateral inhibition produce well-separated secretory cells in such a dynamic tissue? Our theoretical model and experimental measurements of the cellular rearrangement rates indicate that patterning through lateral inhibition alone is unfeasible. Instead, our cell tracking and photoablation experiments point to a different model, which we call 'lateral separation'. This model consists of two ingredients: First and most importantly, secretory neighbors are preferentially separated by the active shrinkage of their shared interface. Secondly, the pattern is stabilized against disruption by the decreased deformability and mobility of secretory cells in response to the stresses exerted by nearby cell extrusions.

Our measurements of interface lengths upon myosin inhibition indicate that increased contractility on the stem-stem interfaces could be the driver the secretory cell separation. This stem cell-driven mechanism fits well with studies that show that it is myosin-based constriction of the apical surface of stem cells, rather than Paneth cells, that generates the curved geometry in the crypt[219,241]. In turn, the observed stability of the interspersed secretory cell pattern upon cell extrusion aligns well with the limited in vivo evidence available[262]. Here, intravital photoablation experiments, where cells (between one and three) in the crypt were ablated, showed that the Paneth-stem checkerboard pattern is restored within two hours by the movement of nearby stem cells into the extrusion site. Using our high-resolution ex vivo data, we can now link this behavior to the high deformability of stem cells and show that Paneth cells often remain separated throughout the extrusion process. The pattern is thus not so much restored during these extrusions but rescued by stem cell deformations and never truly disrupted in the first place.

The role of mechanics in secretory cell pattering also fits a broader picture that is emerging about intestinal cell differentiation, where cells commit early to a specific cell type and only afterwards sort to their proper position[88,272], likely in a manner that relies on cell mechanics. Indeed, Paneth cells have been shown to depend on cell adhesion molecules to stay confined to the stem cell zone[273,274]. Our work now shows that their interspersed patterning in the crypt is mechanically determined as well. It thereby also provides a new perspective on the remarkable stability of the crypt upon stem cell or Paneth cell ablation, where cell mechanics could cause back-migration of TA-cells or other secretory cells respectively to reestablish the proper checkerboard pattern[263,264].

Recently other studies have showed a role for cell movement and mechanics in conjunction with notch signaling in other systems a well[275], ranging from the zebrafish heart[276] to the drosophila eye[277]. Of these systems the patterning of the sensory hair cells in the mouse and chicken inner ear intestine shares the most obvious similarities with the intestine. In both these cases, it seems that surrounding cells drive the separation of the sensory/secretory cells by sharing highly contractile surfaces among themselves[278,279]. Recently, it was shown that in the chicken inner ear this contractility is regulated by myosin and especially its doubly

phosphorylated form[279]. We also establish a role for myosin through our myosin inhibition experiments, but were unfortunately not able to successfully stain for double phosphorylated myosin. Similarly to the intestine, in the mouse inner ear, hair cells are stiffer, rounder, and move less in response to nearby cell extrusions[278]. The mechanical features of 'lateral separation' thus seem to be as conserved between these systems as the lateral inhibition itself.

In this study, we do not examine the link between notch signaling and the 'lateral separation' mechanism. The similarities in cell separation dynamics we observe between the different secretory types in the intestine as well as more general similarities in cell mechanical properties observed between intestinal secretory cells and the hair cells in the inner ear[278], which we note are all cell types that depend on Delta-Notch signaling for their differentiation, do suggest an important role for Notch in mechanical regulation. The link between notch and cell mechanics could be regulated transcriptionally, whereby the cleaved intracellular domain of notch (or the absence thereof) induces the expression of membrane proteins which in turn regulate the tensions on the different cell type interfaces. Alternatively, one could imagine a role for notch directly at the interface, coupling signaling and mechanics much more directly. Interestingly recent studies have shown that the notch trans-membrane domain can indeed locally affect the cytoskeleton cortex[280-282]. Addressing the role of notch signaling could make use of mosaic organoids and tissues where a fraction of the cells constitutively activate intracellular notch signaling, independently of their neighbors, breaking the lateral inhibition mechanism while retaining both notch-positive and negative cell types[283,284]. This would enable disturbing Notch signaling at the membrane to study its mechanical effects without converting cell types. In such mosaic models of the mammary gland and other epithelia, proper pattering is still maintained without functioning lateral inhibition[283,284], suggesting that in these cases too lateral inhibition is complemented with a cell-sorting mechanism.

There is also a striking difference between the intestinal epithelium and sensory epithelia such as the inner ear. In the latter, cell fate is assigned mostly after cells have stopped dividing[285] and after its development is complete it remains completely static. In the intestine there is no such temporal separation of cell division and pattern formation, the whole system is constantly regenerating and rearranging. Even the secretory progenitors divide, further compromising the patterning. While the role of mechanics seems to be to stabilize and refine the checkerboard pattern defined by lateral inhibition in the inner ear[278], it plays the main part in patterning the intestine. The synergy of two mechanisms thus ensures robust patterning across different epithelia: In more static, developmental systems lateral inhibition dominates, while in highly dynamic contexts with frequent cell divisions 'lateral separation' dominates. Besides the intestinal epithelium, the patterning of notch-positive and negative cells in other self-renewing tissues like the lung and mammary gland epithelium, would belong in the latter category. Together lateral inhibition and 'lateral separation' make for an extremely robust catch-all solution for checkerboard-pattern formation, that is more defined by its ends, the pattern, then by its means, cell signaling or mechanics.

6.4 Methods

Organoid culture. For the cell tracking we used mouse intestinal organoids with an H2B-mCherry reporter were used, gifted by N. Sachs and J. Beumer (group of H. Clevers, Hubrecht Institute). For the photo-ablation and 2D organoid experiments we used a E-cadherin-mNeonGreen line gifted by D. Krueger (group of H. Clevers, Hubrecht Institute)[286]. Organoids were grown embedded in membrane extract (BME, Trevigen) in medium consisting of murine recombinant epidermal growth factor (50 ng/ml, Life Technologies), murine recombinant Noggin (100 ng/ml PeproTech), human recombinant R-spondin 1 (500 ng/ml, PeproTech), N-acetylcysteine (1 mM, Sigma-Aldrich), N2 supplement (1×, Life Technologies) and B27 supplement (1×, Life Technologies), GlutaMAX (2 mM, Life Technologies), HEPES (10 mM, Life Technologies) and penicillin–streptomycin (100 U/ml, 100 µg/ml, Life Technologies) in Advanced DMEM/F-12 medium (Life Technologies). Organoids were kept in incubators at 37 °C with 5% CO2. The medium was changed every 2 d or 3d. Each week, organoids were mechanically broken, and the fragments were reseeded.

3D sample preparation. Organoids were broken and seeded 2 days before imaging in four-well chambered coverglass (#1.5 high-performance coverglass) from Cellvis. For the organoids to be within the working distance of the objective and to minimize the required laser power, we let the organoids sink towards the glass by placing the sample on a cold block (~4 °C) for 10 min after seeding. Afterward, the BME gel was allowed to solidify at 37 °C for 20 min before adding medium.

2D organoid culture. We prepared 160 µl gel by mixing 100 µl of lyophilized collagen, 10 µl PBS (10X), 3 µl NaOH (0.5M) and 50 µl BME (Trevigen) in a cold block. Using a wide bore tip, we spread the gel over the bottom of each wells in an 8-well chambered coverglass (#1.5 high-performance coverglass) from Cellvis. We let the gel solidify for at least one hour in an incubator at 37 °C.

Organoids from a 3D culture were broken and digested into single cells by trypsinization (in 1 mL Trypsin-EDTA (0.25%), Thermo-Fisher) for 3 minutes at 36 °C. Afterwards we added 3 ml of cold medium and spun down the cells at 1.8 (x1000) rpm for 5 minutes. The single cells were resuspended in 360 µl of pre-heated medium and then added onto the gel (40 µl per chamber). During the first two days of culture the medium was supplemented with 10 µM Y-27632 (Sigma-Aldrich) to prevent cell death. It was removed from the medium at least a day before imaging.

Microscopy. Imaging was performed on a Nikon A1R MP microscope with a ×40 oil-immersion objective (numerical aperture, 1.30). 30 z slices with a step size of 2 µm were taken per field-of-view every 12 min, with a pixel size of 0.32 μm^2. For the 2D organoids, an image was taken every 20 minutes. For live staining of CD24, APC anti-CD24 (ThermoFischer 17-0242-80) was used at a 1:200 concentration. It was added 1 hour before imaging.

Cell tracking. Cell tracking was done automatically using the OrganoidTracker 2.0 software using the organoid machine learning models described in Betjes et al.[235]. To compute the error rates, we performed the marginalization step with the temperature parameter set to 1.5, matching the original settings. Before any downstream analysis we filtered out all steps in the cell tracks with an estimated error rate above 1%. This left us with a dataset of fragmented cell tracks of high quality, which we could use to measure rates of events like neighbor cell rearrangements and cell division that do not require long tracks, spanning multiple cell cycles, to be detected. For the examples of divisions shown in **Figure 6.2D**, manual tracking was performed.

Quantification of neighbor pair formation and separation rates. We identified neighbor and quantified neighbor relations based on nuclear positions, as outlined in the main text. We first collected the 20 closest neighbors for every cell detection, which we used as list of potential neighbor pairs. For each potential neighbor (*AB)* pair with distance $d_{AB} < 20$ µm, we found the nearest common neighbor, defined as the cell *S* that minimizes the summed distance $d_{AS} + d_{SB}$. We then computed the neighbor score as $\nu = (d_{AS} + d_{SB})/d_{AB}$ and considered all pairs *AB* with a score $\nu > \sqrt{2}$ to be neighbors.

To quantify the rate of neighbor pair formation, we tracked how often a new neighbor pair was formed, measured the neighbor score for a cell pair increasing beyond $\nu = \sqrt{2}$ for at least 1 hour (5 frames). Similarly, we only considered a neighbor pair to have separated, if the neighbor score for a cell pair decreased below $\nu = \sqrt{2}$ for at least 1 hour. Cells that gained or lost contact less than 1 hour before or after division were not taken into account. When it could not be determined if a cell was going to divide or had just divided, due incomplete cell tracks, we also excluded these events. The rates were calculated as $r = (\text{frame rate}) \cdot (\#\ \text{events}) / (\#\ \text{neighbor pairs})$, where depending on the rate measured, events refer formation of a new neighbor pair or separation of an existing pair. Counts of events and neighbor pairs refer to the cumulative number counted at every frame for the duration of the experiment, where events are only counted when they meet the criteria listed above and neighbor pairs were only counted if they could have resulted in a counted even, meaning that the pair would not have been excluded due to divisions or tracking uncertainty, as specified above. This means we only include neighbor pairs, where both cells in the pair were more than an hour away from division. The fraction of number events over number of neighbors is measured is multiplied by the frame rate to get to an event rate per unit time.

Quantification of cell division rate. We identified cell divisions using the division prediction produced by OrganoidTracker 2.0[235]. First, all cells that whose cell tracks split into two daughter cells were considered to be dividing. Secondly, by automatically filtering out potential errors from our tracking data we tend to cut cell tracks short just before their division, as these are among the more difficult events to track. To avoid undercounting divisions we looked at the predicted division probability at the end of all cell tracks. We compared this division probability against the chance of the cell dying or having a missed

detection in the next frame, the other major source of tracks being lost. If the division probability at the end of the track was thus higher than p=0.01, (estimated probability of cell death or a missed detection) we counted it as a division, as we consider it the most likely explanation for the track being lost. The division rate (per cell per unit time) was then calculated as the ratio of the total number of division events and the total number of cells detections counted over every frame and multiplying this by the framerate, $r = (frame\ rate) \cdot (\#\ divisions) / (\#\ cell\ detections)$.

Quantification of cell type clustering. For cell type specific analyses, we reused the data from Zheng et al.[88]. To summarize, the data consists of 5 organoids tracked for upwards of 60 hours, after which antibody staining was performed to identify most intestinal cell types: stem cells, Paneth cell, mucus-producing cells (MPCs) and enteroendocrine cells (EECs). Using the lineage information, we can use the endpoint measurements to infer the cell fate identity of practically all cells during the experiment itself[88].

We computed the tendency of two cell types A and B to cluster together as $\frac{[AB]}{[A][B]}$, where $[AB]$ is fraction of AB neighbor pairs, as determined by their neighbor score, and the product $[A][B]$ indicates the expected fraction of AB neighbor pairs if cells were randomly distributed, given simply by the product of the concentrations $[A]$ and $[B]$ of the individual cell types. To calculate the clustering of secretory cells in the stem cell zone, we only considered a subset of cells for which at least half of their neighbors belonged to a type commonly found in the stem cell zone (stem cells, Paneth cells, MPCs, EECs).

Temporal analysis of cell type-specific cell neighbor pair separation. For every neighbor pair we measured the time between when two cells came into contact ($v>\sqrt{2}$ for at least 1 hour) and lost contact again ($v<\sqrt{2}$ for at least 1 hour). This gave us a series of lifetimes for all neighbor pairs in the data. For two given cells, this might result in multiple datapoints as they might gain contact again after losing it. We then used survival analysis (R-package *survival*) to analyze the lifetimes of neighbor pairs in a way that can deal with observations being cut short by reasons not related to cells moving apart, which could otherwise lead to overestimates of the rate at which neighbor pairs are remodeled. We computed the fraction of neighbor pairs remaining in contact over time, with $t=0$ defined as the moment the cells came into contact either because one cell has just been born or because they came together during a T1-transition. When either of the cells was lost to tracking, we considered this a censuring event. A cell division or cell death of either of the cells was treated the same. We used the lineage information and endpoint staining to infer the cell type for the cells making up the neighbor pairs, which allowed us to stratify the survival analysis according to the cell types involved.

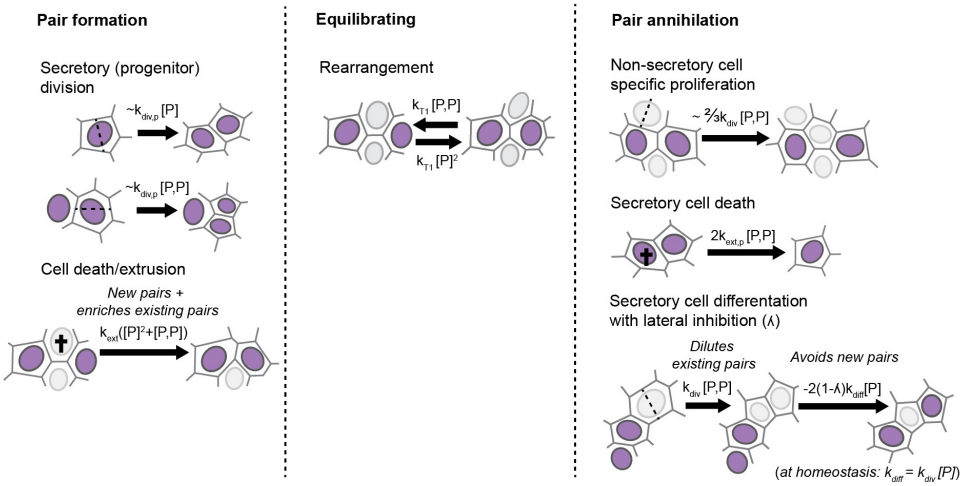
Photoablation. Laser ablation was performed with an 800 nm 2-photon laser at 100% laser power for 8 msec. For laser ablation, we searched for a configuration of two secretory cells in the crypt, as identified by staining with a conjuged CD24-antibody (ThermoFischer 17-

0242-80), that were separated by one cell in between. If the configuration was such that the two secretory cells were connected by a single interface between two adjacent non-secretory cells, as indicated in Figure 6.4b in the main text, we aimed the 2-photon laser at this interface, typically causing these two non-secretory cells to extrude out of the epithelium. We located the focus point of the laser on the basal side of the interface in between the nuclei. Cells were only targeted when their basal side was facing the objective. Less frequently, two secretory cells were separated by a single non-secretory cell in between and hence no such interface could be located. In this case, we aimed the 2-photon laser at the basal side of the single non-secretory cell separating the secretory cells. Whether laser ablation resulted in successful extrusion was typically predicted by the appearance in the transmitted light microscopy channel of a small bubble forming at the site on which the laser pulse was focused. We performed 10-20 ablations in the same tissue, within a 20-minute time window, with a fraction of those ablation-induced extrusions resulting in secretory cells becoming neighbors. The data for photoablation was generated from 5 different experiments. Afterwards, we imaged overnight for 14 to 16 hours.

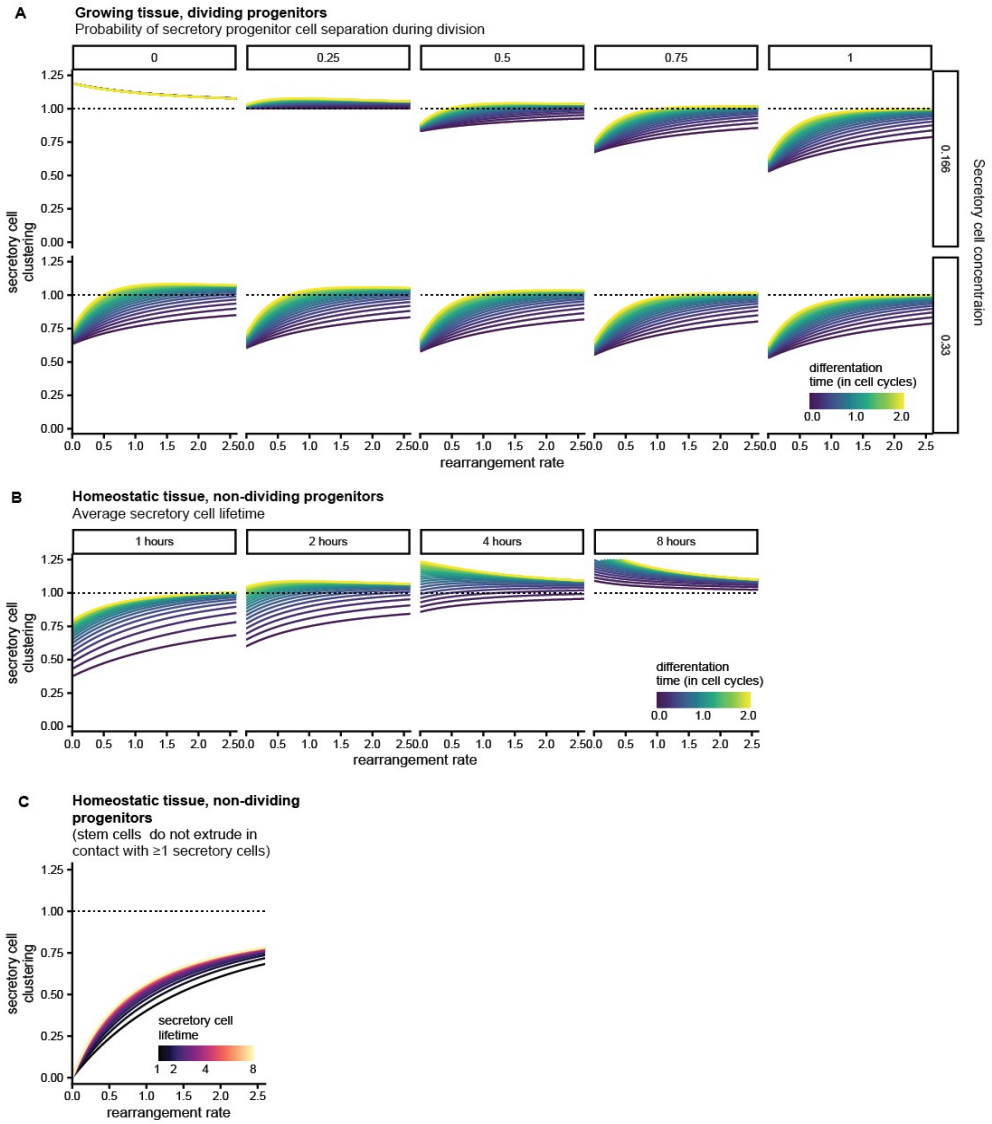
Quantification of interface length. The interface lengths of each Paneth-Paneth pair and the control pairs were measured with OrganoidTracker[235]. The interfaces of the cell membranes of each pair were marked manually at the tricellular junctions at the interface ends. The distance between the points was calculated and defined as the interface length. For photoablation experiments, which were performed in 3D culture, all interface lengths were measured in a single z-slice on the basal side, at the height where the nuclei were located. This facilitated more exact measurements than at the apical side, where the tissue was highly curved. For 2D culture, we analyzed lateral interface lengths as close to the apical membrane where a defined E-cadherin signal was still visible, as E-cadherin does not localize on apical membranes. While the tissue relaxes during the experiment due to the Blebbistatin treatment, the height of the apical interface might change and the z-plane used to quantify interface lengths was changed accordingly.

Palbociclib and Blebbistatin treatment. We subjected the 3D organoids to Palbociclib (Sigma-Aldrich) at a final concentration of 10 μM at the start of imaging, after which organoids were then imaged for 2 days. Three Palbociclib-treated organoids and three control organoids with limited cell death were chosen from two different experimental days. 2D organoids were treated with 10 μM Blebbistatin (Sigma-Aldrich) just before imaging, after which the organoids were imaged for at least 1 day. The examples shown are picked from two separate experiments.

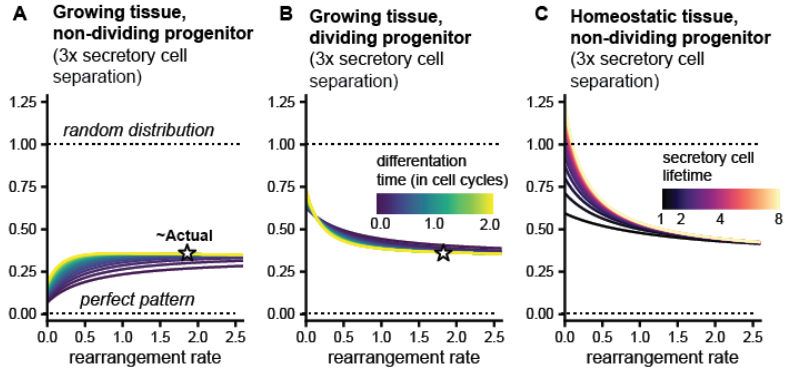
6.5 Supplementary Figures



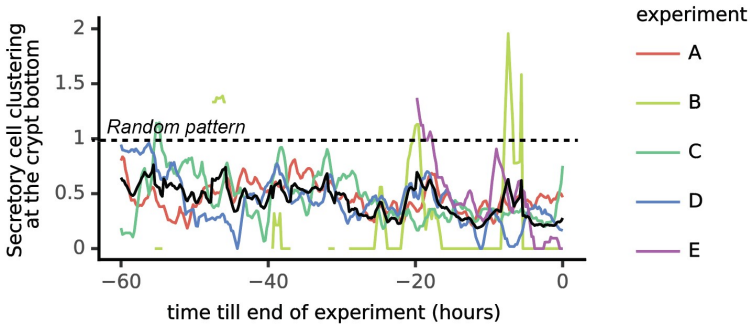
Supplemental figure 6.1: Transitions in the model at the microscale. Left column: mechanisms that cause formation of secretory cell neighbor pairs, specifically, secretory cell divisions and extrusions of non-secretory cells. The latter act both by bringing (potentially secretory) cells together and by enriching the relative concentration secretory neighbor pairs. Secretory cell neighbor pairs are measured by the concentration of secretory cell pairs $[P,P]$, while the concentration of secretory cells is given by $[P]$. Middle column: randomizing mechanisms that equilibrate the secretory pair concentration. This is done by T1-transitions that can both bring together and separate secretory neighbors. Right column: mechanisms which induce separation of secretory pair, leading to a decrease in $[P,P]$. Proliferation of nearby cells tends to force apart non-proliferative secretory cells. Secretory cell death also removes secretory pairs from the system. Differentiation through lateral inhibition lowers the concentration of secretory pairs, while expansion of the non-secretory pool first dilutes the secretory pair concentration. To maintain secretory cell homeostasis, non-secretory cells must differentiate but lateral inhibition insures they do not create secretory pairs in the process. We denote the lateral inhibition efficiency with λ and define it as the efficiency at which formation of new secretory pairs through secretory differentiation is avoided.



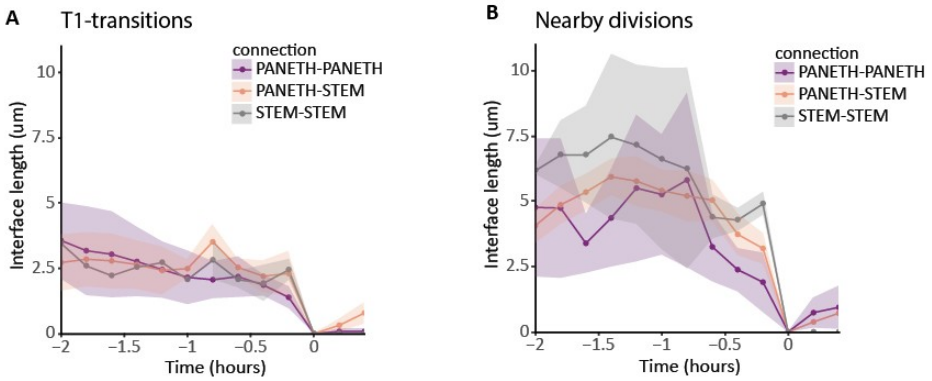
Supplemental figure 6.2: Alternative model parameterizations. A) Model of a growing system with dividing progenitors that have an increasing probability of separating during division (from left to right). Secretory cell clustering (φ) is characterized across different differentiation times (in color) and different secretory cell concentrations (top to bottom). At low secretory cell concentrations (0.166) separation of sisters helps patterning by avoiding the creation of secretory sister neighbors. At high concentrations (0.33) the effect is negligible because the separation process will create new secretory neighbors with secretory cells further away (see **Appendix**). Dotted line indicates a randomized pattern. **B)** Secretory cell clustering in homeostatic tissues at different secretory cell lifetimes (left to right) and across different differentiation times (in color). Patterning (φ<0.75) only emerges at low differentiation times and short secretory lifetimes. **C)** Model for homeostatic tissue where non-secretory cells can only leave the system when they are not touching any secretory cells. This avoids the creation of new secretory pairs during their extrusion and improves patterning (compare with **figure 6.1F**). Still, we only see the measured patterning values (φ~0.3) at low rearrangement rates (less than 0.5 times the cell division rate).



Supplementary figure 6.3: Model with faster secretory cell separation. Secretory cell clustering when secretory pairs move apart at three times the rate compared to the overall rearrangement rate in the system. The same systems are shown as in in Figure 6.1D,E,F. Star denotes the measured secretory cell clustering and rearrangement rate values.



Supplementary figure 6.4: Secretory cell patterning when including early progenitors. Quantification of secretory cell clustering in the stem cell zone during the timelapse. At the final timepoint secretory cells are identified using mature markers. At earlier timepoints we can leverage the lineage information to identify early secretory progenitors (that not yet express markers) and include them in the pattern quantification. Including early progenitors does not improve secretory patterning. The clustering parameter is averaged over a hour long window, black line indicates mean across the experiments

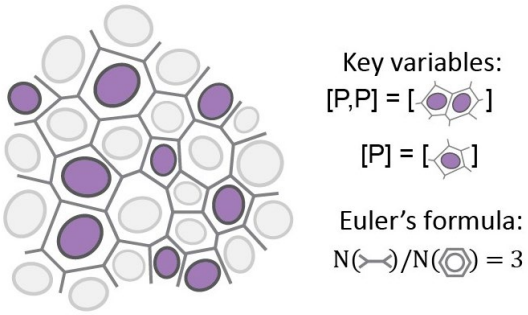


Supplementary figure 6.5: Different modes of cell separation. A) The interface lengths of neighbor pairs undergoing T1-transitions aligned at their moment of separation. Interface length shows a gradual decline towards zero and no difference between cell types can be seen. Shaded area denotes standard error of measurement. **B)** The same as in A) but for neighbors separated by nearby divisions. All types of neighbor pairs exhibit a rapid decreasing interface length shortly before moving apart but have large interfaces before that.

6.6 Appendix: Model description

6.6.1 Modelling of lateral inhibition in the presence of cell movement

Our goal is to model the steady state concentration of neighboring secretory cell pairs (denoted $[P,P]$) at a constant concentration of secretory cells (denoted $[P]$). The total number of cells is allowed to change to account for tissue growth. Our model is zero-dimensional and only explicitly follows the dynamics of the concentration of the secretory neighbor pairs, but is underpinned by a 2D picture of the intestinal epithelium to derive the various rates of secretory pair formation and separation processes **(Appendix figure 6.1)**.



Appendix figure 6.1: Dynamics of neighbor pair concentration from 2D tissues principles. Cartoon; a confluent 2D tissue with secretory (purple) and non-secretory cells (grey). The key variables in our model are the concentration of secretory neighbor pairs ($[P,P]$) and the concentration of secretory cells ($[P]$). These variables have no spatial component and are not linked to a specific underlying cell configuration, but the general geometric principles of 2D tissues do inform the rates governing the neighbor pair concentrations through Euler's law, which specifies a mathematical relationship between the numbers of interfaces, junctions, and cells for any plane-filling tiling of cells[270]. In particular, when considering a rate k_{cell} for events occurring at the cell-level, such as division, the rate of the same process, but then for a neighbor cell pair, is given by $\frac{1}{3}k_{cell}$, with the prefactor of $\frac{1}{3}$ implementing that, according to Euler's formula, in 2D epithelia there are three times as many neighbor pairs (faces) as cells in the system.

While the model takes into account spatial structure at the level of neighbor pairs, it does not keep track of higher order spatial structure. However, rearrangements bring cells into contact with more distant cells than direct neighbors, with the identity of these cells important for the type of new neighbor pair formed. Here, we make use of a key assumption, namely that beyond neighbor cell pairs the system is well-mixed, meaning that during a rearrangement the probability of a more distant cell being a secretory cell is simply given by the overall secretory cell concentration $[P]$. In the final section we discuss this assumption in more detail. Furthermore, we only consider two types of cells, secretory cells (at concentration $[P]$) and non-secretory stem cells or transit-amplifying cells (at concentration $1-[P]$), depending on which part of the crypt we model. For the systems we aim to describe the stem cells or TA cells can differentiate into secretory cells, but these cannot dedifferentiate. Furthermore, we allow the secretory cells to have a different (or a zero) division rate and a different extrusion

rate than the overall population. Otherwise, our assumptions are that the secretory cells are identical in all respects to the overall population.

6.6.2 Deriving rates for secretory pair formation and disappearance

We first note that the change in secretory neighbor pair concentration is determined by the rate of change in the number of secretory pairs ($N_{P,P}$) and the rate of change in total neighbor pair number (N):

$$\frac{d[P,P]}{dt} = \frac{d(\frac{N_{P,P}}{N})}{dt} = \frac{1}{N}\frac{dN_{P,P}}{dt} - \frac{N_{P,P}}{N^2}\frac{dN}{dt} = \frac{1}{N}\frac{dN_{P,P}}{dt} - [P,P]\frac{1}{N}\frac{dN}{dt}$$

At steady state secretory pair concentration, this means that the rate of secretory neighbor creation matches the tissue growth rate:

$$\frac{1}{N}\frac{dN_{P,P}}{dt} = [P,P]\frac{1}{N}\frac{dN}{dt}$$

The growth rate is determined by the average division (k_{div}) and extrusion (or death) rate (k_{ext}) of cells in the system:

$$\frac{dN}{dt} = N(k_{div} - k_{ext})$$

Such that the rate of change of secretory pairs ($\frac{1}{N}\frac{dN_{P,P}}{dt}$) has to match:

$$\frac{1}{N}\frac{dN_{P,P}}{dt} = [P,P](k_{div} - k_{ext})$$

To determine the rate at which the number of secretory pairs changes ($\frac{1}{N}\frac{dN_{P,P}}{dt}$), we will identify the scenarios by which secretory pairs can emerge and disappear in the following paragraphs. Before this, we note that the secretory cell concentration is given by $[P] = \frac{n_P}{n}$, with n referring to the number of cells rather than of cell neighbor pairs. The change in the number of secretory cells is then given by:

$$\frac{dn_P}{dt} = \left(k_{div,P} - k_{ext,P}\right)[P] \cdot n + k_{diff}(1 - [P]) \cdot n$$

where k_{diff} is the rate of differentiation of secretory cells from stem or TA cells and where we allow secretory cells to divide and extrude at rates $k_{div,P}$ and $k_{ext,P}$ that can differ from that of the overall population rate k_{div} and k_{ext}.

We will now move to our description of secretory pair emergence and disappearance.

T1-transitions.

The T1-transistion rate (k_{T1}) is the rate at which a neighbor pair moves apart to let a new pair form in between. The rate at which secretory pairs disappear is then simply given by this rate multiplied by the fraction of secretory pairs ($k_{T1}[P,P]$). The chance of a new secretory pair

forming is given by the chance that both cells of the new pair are secretory cells ($[P]$), which under the assumption that the system is well-mixed, is given by $k_{T1}[P]^2$. The rate at which the concentration of secretory pairs changes due to T1-transitions is then given by:

$$k_{T1}[P]^2 - k_{T1}[P,P]$$

Secretory cell differentiation with lateral inhibition.

Secretory differentiation results in a secretory pair only if a non-secretory cell differentiates (with rate $k_{diff}(1-[P])$ in events per unit time) when its neighbor is a secretory cell (with probability $[P]$). Because the new secretory cell has on average six neighbors, the rate of secretory neighbor differentiation without lateral inhibition is given by:

$$6 \cdot \frac{1}{3} k_{diff}(1-[P])[P] = 2 k_{diff}(1-[P])[P]$$

Here, the factor $\frac{1}{3}$ reflects a conversion of a cell-based rate to a rate for neighbor pairs, following Euler's formula to account for the structure of 2D epithelia (**Appendix figure 6.1**). In our model, lateral inhibition acts to reduce the number of newly emerging secretory pairs. We define its efficiency (λ) as the fraction of such pairs it can avoid:

$$2(1-\lambda)k_{diff}(1-[P])[P]$$

with the highest possible efficiency given by λ =1. In our model, we assume that Notch signaling fully inhibits commitment to secretory cell fate in neighbor cells. However, a reduced efficiency of lateral inhibition, $\lambda < 1$, can still arise because rearrangements occurring after commitment can bring these cells into contact with existing secretory cell, which we discuss in more detail in the next section. When we enforce that $\frac{d[P]}{dt} = 0$, i.e. the secretory cell concentration $[P]$ remains constant in time, k_{diff} is not an independent parameter but instead determined by the division and extrusion rates of the overall population (k_{div}, k_{ext}) and the secretory cell population ($k_{div,P}, k_{ext,P}$). Indeed, using the chain rule one arrives at (with n being the number of cells in the system):

$$\frac{1}{n}\frac{dn_P}{dt} = [P]\frac{1}{n}\frac{dn}{dt}$$

$$k_{diff}(1-[P]) + [P]\left(k_{div,P} - k_{ext,P}\right) = [P](k_{div} - k_{ext})$$

$$k_{diff} = (k_{div} - k_{div,p} - k_{ext} + k_{ext,p})\frac{[P]}{(1-[P])}$$

which gives us for the contribution of differentiation to secretory pair creation:

$$2(1-\lambda)(k_{div} - k_{div,p} - k_{ext} + k_{ext,p})[P]^2$$

Secretory cell division.

An isolated secretory cell can give rise to a secretory cell neighbor pair by dividing, with division rate $k_{div,p}[P]$ (events per unit time per cell), corresponding to a rate per neighbor pair of $\frac{1}{3}k_{div,p}[P]$ following Euler's formula. We first consider the simplest scenario where sister cells remain adjacent following the division (corresponding to **figure 6.2D,iii** in the main text). Such a division will give rise to three new cell pairs (**Appendix Figure 6.2**). One cell pair corresponds to the two secretory sisters and will hence always add one secretory neighbor pair by definition. The two remaining cell pairs concern neighbor cells linked to the mother cell via a cell interface bisected by the mother cell's division plane. Hence, for each of these neighbor cells, the probability that they form a secretory neighbor pair with the secretory daughter cell is given by the probability that the neighbor cell is also a secretory cell, which is given by $[P,P]/[P]$. Hence, the contribution of these three new pairs to the concentration of secretory neighbor pairs is given by:

$$\frac{1}{3}k_{div,p}[P]\left(1+2\frac{[P,P]}{[P]}\right)=\frac{1}{3}\,k_{div,p}[P]+\frac{2}{3}k_{div,p}[P,P]$$

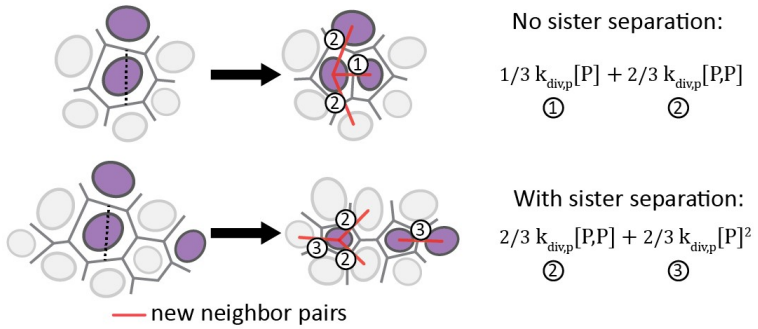
The first term here represents the rate of secretory sister neighbor formation while the second term denotes the rate at which secretory neighbor pairs are created with one newly divided and one existing secretory cell.

For completeness, we show here that with the same approach we can also incorporate different cell division patterns. Here, we discuss the division pattern where the secretory sister cells separate during division and are both displaced relative to the mother cell's position (corresponding to **figure 6.2D,i** in the main text). Again, tracking new cell pairs formed by the division, in this case followed by sister displacement, we obtain four new neighbor pairs that might contribute to additional secretory cell neighbor pairs (**Appendix figure 6.2**). Two pairs involve neighboring cells already in contact with the mother cell and hence, as above, the probability that these neighbors were secretory cells is given by $[P,P]/[P]$. The remaining two pairs are each formed by one sister gaining contact when a cell that was not a neighbor of the mother cell. Here, the probability that this new cell is secretory and, hence, that they form a secretory neighbor pair with one of the sister cells is given by $[P]$. Overall, the resulting rate of formation of secretory neighbor pairs is thus given by:

$$\frac{1}{3}k_{div,p}[P]\left(2\frac{[P,P]}{[P]}+2[P]\right)=\frac{2}{3}k_{div,p}[P,P]+\frac{2}{3}\,k_{div,p}[P]^2$$

We can combine this division pattern and the pattern above, by introducing a parameter α that denotes the probability of the sisters remaining together after division, yielding the following secretory neighbor pair formation rate:

$$\frac{1}{3}\alpha\,k_{div,p}[P]+\frac{2}{3}k_{div,p}[P,P]+\frac{2}{3}(1-\alpha)k_{div,p}[P]^2$$



Appendix figure 6.2: Illustration of the role of secretory cell division in secretory pair creation. Top; a secretory cell (purple) divides with sister cells remaining adjacent at the mother cell's original position. This division pattern given rise to two classes of new neighbor cell pairs (indicated by the redline): the neighbor pair formed by the two sisters (labeled 1); and two neighbor pairs formed between one sister cell and the cell neighboring that sister through the cell interface bisected by the division plane (labeled 2). We note that in the cartoon we assigned these latter new pairs arbitrarily to one sister, while the equivalent pairs for the other sister are assumed the same as the pairs that existed for the mother cell and, hence, do not contribute to the creation of new neighbor pairs. These newly formed neighbors pair classes increase the concentration of secretory neighbor pairs with the rate shown next to the cartoon. Bottom; division of a secretory cell leads to secretory daughters separating and displaced away from the mother cell position. Hence, no neighbor pair of sister cells will emerge (corresponding to class 1 in top panel). New neighbor pairs (labelled 2) include two new pairs with cells that neighbored the mother cell, but also two new pairs with cells not previously in contact with the mother (labelled 3). The contribution of these new neighbor pairs to the rate at which new secretory neighbor pairs are formed is shown next to the cartoon

Secretory cell death.

We note, as above, that in the intestine cell death and cell extrusion are equivalent. Secretory cell death (k_{ext}) simply removes secretory pairs at twice the death rate, as there are two secretory cells in the pair, yielding the following rate:

$$-2k_{ext,P}[P,P]$$

New neighbors created by cell death.

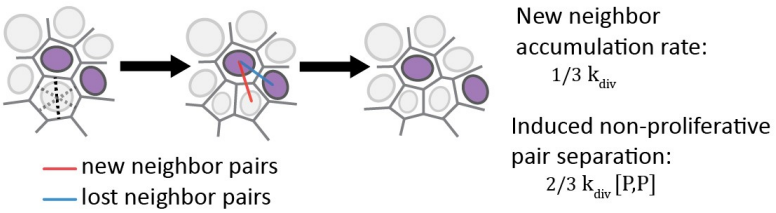
Cell extrusions forces neighboring cells together into new pairings. On average six neighbor pairs are lost during a single extrusion, which for pairs occurs at rate $\frac{1}{3}k_{ext}$. This loss must be compensated by the creation of three new cell pairs to comply with Euler's formula, which requires that the ratio of neighbor pairs to cells equals three. These have a probability of being secretory pairs, assuming a well-mixed system, that depends on the square of the secretory cell concentration $[P]^2$. This leads to the following rate:

$$3 \cdot \frac{1}{3}k_{ext}[P]^2 = k_{ext}[P]^2$$

Cell division effects on adjacent non-proliferative cells.

While the model allows for divisions of secretory cells, in practice the secretory division rate $k_{div,P}$ will at least be much smaller than that of non-secretory cells and in some scenarios even be set to zero. For such non-proliferating cells, divisions of neighboring cells increase the number of neighbors (**Appendix figure 6.3**). When non-proliferating cells are surrounded by proliferating cells, they would therefore accumulate neighbors without bounds, unless rearrangements occur that bring back the number of neighbors to the average of six found in 2D tissues. We therefore assume that non-proliferating cells undergo an increased rate of T1 transitions that separate them from their neighbors. Non-proliferating cells gain new neighbor pairings from divisions at a rate of $6 \cdot \frac{1}{3} k_{div}$, as they have on average six neighbors which have a chance a $\frac{1}{3}$ to orient their division plane such that the cell gains a new neighbor (**Appendix figure 6.3**). This must be compensated for by a similar increase in the T1 rate. Because the cell again has on average six neighbors, this is reflected by an increase per neighbor pair of $\frac{1}{3} k_{div}$. If secretory cells do proliferate, but more slowly than the rest of the tissue ($k_{div,P} < k_{div}$), we must adjust this rate by the cell division rate. This is because daughter cells on average lose one neighbor compared to their mother. Cell divisions thereby compensate for the neighbors gained during the cell division. This is also the reason why cells proliferating at the same rate as the overall tissue (secretory cells if $k_{div,P} = k_{div}$), do not have to undergo extra T1-transitions[270]. For a secretory pairs (at concentration $[P,P]$), both secretory cells have altered T1-transition rates, this combines to the following contribution to the secretory pair formation rate:

$$-2 \cdot \frac{1}{3}\left(k_{div} - k_{div,p}\right)[P,P]$$



Appendix figure 6.3: Illustration of the role of cell divisions on nearby secretory cells: A non-proliferating secretory cell (purple) can gain neighbors when a neighbor divides. If this actually happens for a given non-proliferating cell depends on the orientation of the division plane of the dividing cell, as given by the factor 1/3, as a dividing cell will create a new neighbor for two out of the average six cells surrounding it. To avoid an ever-increasing number of neighbors, the non-proliferating cells need to lose their neighbors through T1-transitions in the final step. This leads to increased separation rates of pairs of non-proliferating cells

6.6.3 Full model

By combining all the components described above, we obtain a single equation for the steady state level of the secretory pair fraction $[P,P]$. For simplicity, we set the probability of sisters remaining neighbors during division to α=1. This is supported by our model result that for the secretory cell concentration we use, $[P]=\frac{1}{3}$, the effect of secretory cell separation during division is negligible (**Supplementary figure 6.2A**). This results in:

$$k_{T1}[P]^2 - k_{T1}[P,P] - \tfrac{2}{3}\left(k_{div}-k_{div,P}\right)[P,P] \qquad \textit{(cell rearrangement)}$$

$$+\tfrac{1}{3}k_{div,P}[P] + \tfrac{2}{3}k_{div,P}[P,P] \qquad \textit{(secretory cell division)}$$

$$+k_{ext}[P]^2 - 2k_{ext,P}[P,P] \qquad \textit{(cell extrusion)}$$

$$+2\,(1-\lambda)(k_{div}-k_{div,P}-k_{ext}+k_{ext,P})[P]^2 \qquad \textit{(differentiation, lateral inhibition)}$$

$$=(k_{div}-k_{ext})[P,P] \qquad \textit{(growth of the system)}$$

Our patterning parameter φ at equilibrium can then simply be calculated as:

$$\varphi=\frac{[P,P]}{[P]^2}=\frac{k_{T1}+k_{ext}+\frac{1}{3}k_{div,p}/[P]\ +2\,(1-\lambda)(k_{div}-k_{div,P}-k_{ext}+k_{ext,P})}{k_{T1}+\frac{2}{3}\left(k_{div}-k_{div,P}\right)+\frac{2}{3}k_{div,p}+\ 2k_{ext,P}+(k_{div}-k_{ext})}$$

In a scenario where the secretory cells do not divide ($k_{div,p}=0$) this simplifies to:

$$\varphi=\frac{[P,P]}{[P]^2}=\frac{k_{T1}+k_{ext}\ +2\,(1-\lambda)(k_{div}-k_{ext}+k_{ext,p})}{k_{T1}+\frac{2}{3}k_{div}+\ 2k_{ext,P}+(k_{div}-k_{ext})}$$

In a tissue where the secretory cell extrusions occur at a similar rate as for the overall population ($k_{ext,P}=k_{ext}$) we get:

$$\varphi=\frac{[P,P]}{[P]^2}=\frac{k_{T1}+k_{ext}\ +2\,(1-\lambda)k_{div}}{k_{T1}+\frac{5}{3}k_{div}+\ k_{ext}}$$

This is the scenario plotted in **Figure 6.1D** and **Figure 6.1E,** using a non-dividing and dividing secretory progenitor (see next section) respectively. Proper interspersed patterning of secretory cells implies φ≈0. We can then readily see that patterning only occurs at efficient lateral inhibition ($\lambda\approx 1$) and low rearrangement and cell extrusion rates ($k_{T1}+k_{ext}\ll k_{div}$).

If the tissue does not grow, but cell divisions are compensated by cell extrusions ($k_{div}=k_{ext}$) we get the following:

$$\varphi=\frac{[P,P]}{[P]^2}=\frac{k_{T1}+k_{div}\ +2\,(1-\lambda)k_{ext,P}}{k_{T1}+\frac{2}{3}k_{div}+\ 2k_{ext,P}}$$

Here clustering can only occur at short secretory lifetimes (high $k_{ext,P}$), in addition to requiring efficient lateral inhibition and low rearrangement rates. This is the scenario plotted in **Figure 6.1F.**

6.6.4 Computing lateral inhibition efficiency

The definition of the lateral inhibition efficiency further above suggests it is an independent parameter. However, it is directly linked to cell rearrangements, as we discuss here. We assume that the only cells that can commit to secretory fate are stem or TA cells that are not a neighbor pair with a secretory cell, due to Notch signaling from the secretory neighbor. However, cells do not immediately become a secretory cell after committing to that fate. Instead, they need some time to differentiate before they become recognizable as secretory cells of which we can quantify the pattern. In that time, these secretory progenitor cells can become neighbors of a secretory cell due to rearrangements. As we further assume that these secretory progenitors no longer respond to Notch signaling, this will result in a fraction of secretory cells emerging from a secretory progenitor state next to an existing secretory cell, with the probability of this outcome increasing with higher rearrangement rates and differentiation times.

We can estimate the impact of differentiation of committed secretory progenitors as follows. We assume that all newly emerging secretory cells spend a fixed time T as a committed secretory progenitor. The lateral inhibition efficiency then decreases with the probability $[P]_{pro}$ that a secretory progenitor cell P_{pro} neighbors a secretory cell $[P]$, given by $[P]_{pro} = [P, P_{pro}]/[P_{pro}]$. The lateral inhibition efficiency is then related to this probability evaluated at the time T, divided by the expected fraction of secretory neighbors if the progenitors were randomly distributed, $[P]$, by the following expression:

$$(1-\lambda) = \frac{[P]_{pro}(T)}{[P]}$$

$$\lambda = 1 - \frac{[P]_{pro}(T)}{[P]}$$

Lateral inhibition efficiency is thus optimal if the progenitor neighbors zero secretory cells at the end of differentiation ($\lambda \approx 1$) and no effective lateral inhibition occurs if the local concentration of secretory cells neighboring the secretory progenitor matches the overall secretory concentration.

We model the impact of rearrangements on secretory progenitors as follows. Secretory progenitors can gain or lose secretory neighbors through T1 transitions and nearby cell deaths or divisions. We exclude the possibility of secretory progenitors differentiating next to another secretory progenitor, as this would be avoided through lateral Notch signaling. The change in its neighbor composition is then given by:

$$\frac{d[P]_{pro}(t)}{dt} = k_{T1}[P] - k_{T1}[P]_{pro}(t) - \frac{2}{3}k_{div}[P]_{pro}(t) + k_{ext}[P] - k_{ext,p}[P]_{pro}(t)$$

The above formula assumes no secretory cell divisions,$k_{div,P} = 0$. Because the local concentration of secretory cells around the progenitor takes the form of a conditional probability (what is the chance of finding a secretory cell given that its neighbor is a

progenitor cell, the rates depend not on the squares of the secretory cell concentration ($[P]^2$) but linearly on the secretory cell concentrations.

We can solve this equation in time, to find its value at the differentiation time T. This requires specifying an initial condition. Here, we assume that at the start of the differentiation, the secretory progenitor has zero secretory cell neighbors , $[P]_{pro}(0) = 0$, reflecting that secretory progenitor can only arise from non-secretory cells that do not neighbor a secretory cell, due to the inhibiting effect of Notch signaling. Therefore, this differential equation solves to:

$$[P]_{pro}(t) = [P]_{pro,eq}(1 - e^{-k_{tot}\,t})$$

With k_{tot}:

$$k_{tot} = k_{T1} + \frac{2}{3}k_{div} + k_{ext,p}$$

And the equilibrium:

$$[P]_{pro,eq} = \frac{k_{T1} + k_{ext}}{k_{T1} + \frac{2}{3}k_{div} + k_{ext,p}}[P]$$

The lateral inhibition efficiency (λ) is then given by:

$$(1 - \lambda) = \frac{k_{T1} + k_{ext}}{k_{T1} + \frac{2}{3}k_{div} + k_{ext,p}}(1 - e^{-k_{tot}\,T})$$

This formula for lateral inhibition efficiency is used in the scenario's shown in **Figure 6.1D** and **Figure 6.1F,** where the progenitor is not dividing. In order to have efficient lateral inhibition the differentiation time has to be fast compared to the rearrangement and division rate ($1/T \gg k_{T1} + \frac{2}{3}k_{div} + k_{ext,p}$) or alternatively the division rate has to dominate over the rearrangement rate ($k_{div} \gg k_{T1}$). In the latter case nearby non-secretory divisions tend to push the non-proliferative secretory progenitor cells away from its non-proliferative secretory neighbors (see *Cell division effects on non-proliferative cells*).

If the differentiation process starts after a final symmetric division (shown in **Figure 6.1E**), we have to adjust this description to take into account the fact that the initial state does not have zero secretory neighbors. The average concentration of secretory neighbors at the start is $\frac{1}{6}$ if the sisters remain next to each other under the assumption that the progenitor cell has one average six neighbors one of which is its sister. For completeness we again also consider cases where sister cells move apart (probability of α). In that case , we expect the daughter to connect to at least one random extra cell (**Appendix figure 6.4**), so the mean number of secretory neighbors becomes $\frac{1}{6}[P]$. This gives the following initial state:

$$[P]_{pro,0} = \frac{1}{6}\alpha + \frac{1}{6}(1 - \alpha)[P]$$

The lateral inhibition efficiency is then described by:

$$(1-\lambda) = \frac{k_{T1} + k_{ext}}{k_{T1} + \frac{2}{3}k_{div} + k_{ext,p}}(1 - e^{-k_{tot}\,T}) + \frac{1}{6}\left(\frac{\alpha}{[P]} + (1-\alpha)\right)e^{-k_{tot}\,T}$$

6.6.5 Key assumptions

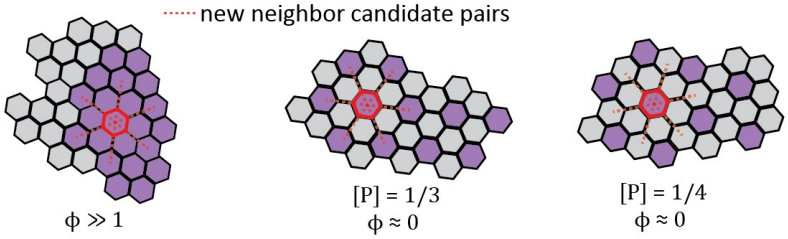
Our most important assumption is that interaction of secretory cells and secretory pairs with the surrounding cells are modeled with a mean-field approximation. Concretely, this means that the fraction of secretory cells at a distance of one cell away from another secretory is taken to be equal to the average concentration of secretory cells. The probability of two secretory cells coming together in the model is thus not influenced by any high-order structure.

In the presence of high order structure, like secretory cell clustering, the local secretory cell concentration around a secretory does not match the average concentration (**Appendix figure 6.4**). This means that the chance of creating new secretory pairs through T1-transitions or cell extrusions occur is higher than our model would predict. Our approximation thus underestimates the secretory pair formation rate (by at most a factor $[P]$). Our model would therefore underestimate the amount of induced extra secretory clustering when the secretory cells cluster ($\varphi > 1$). As we are interested in the opposite of secretory cell clustering ($\varphi < 1$), this thus poses no problem. Furthermore, we are interested in deducing the limits of lateral inhibition in patterning cells, so the fact the model might overestimate the degree of patterning (lower φ) does not weaken our conclusions.

Some highly ordered states without clustering might similarly lead to underestimation of the secretory pair formation rate. At very high concentrations of secretory cells (~33%) and almost perfect patterning ($\varphi = 0$) a hexagonal crystalline structure can similarly cause secretory cells to have more often (or even always) a secretory cell at a distance of one cell away (**Appendix figure 6.4**). Similarly, this could cause an underestimation of some of the secretory pair forming mechanisms by a factor of at most $[P]$. Our steady state solution near $\varphi = 0$ would then become:

$$\varphi = \frac{\frac{1}{[P]}T_1 + k_{ext} + 2\,(1-\lambda)k_{div}}{T_1 + \frac{5}{3}k_{div} + k_{ext}}$$

We note that this effect only occurs in situations where the tissue is already clearly capable of patterning itself and again that we overestimate the patterning capabilities of lateral inhibition. Furthermore, in intestinal organoids we do not see the crystalline patterns that would give rise to this effect (see for instance **Figure 6.3** in the main text).



Appendix figure 6.4: Limitations of the mean-field approximation. Left: for clustering of secretory cells ($\varphi > 1$), a T1-transition involving a secretory cell will almost always lead to a new secretory cell pair. Middle: In the particular case of a perfect hexagonal lattice with exactly a third of the cells secretory, all T1 transitions involving a secretory cell create a new secretory cell pair. Right: An example with a slightly lower secretory cell concentration shows that this is not a general property of configurations with highly dispersed secretory cells ($\varphi = 0$).

7 Discussion

In this thesis we used live imaging to track cells in organoids to study intestinal cell differentiation. We have specifically focused on the relation between spatial position and cell fate as live imaging with its high spatial and temporal resolution is ideally suited to study their interplay. The main results of this thesis are both methodological and biological in nature, where technical developments in our cell tracking and imaging were used directly to make progress in understanding biology.

In **chapter 2,** we review the state-of-the-art at the start of this thesis work, by looking at examples of cell tracking in a developmental biology context. Here, we saw that cell tracking was often limited to small systems and required extensive human curation. In **chapter 3**, we introduce a conceptually new approach to cell tracking that greatly reduces the need for human curation and even enables, for the first time, fully automated analysis of key lineage features like cell cycle length and differentiation rate. We do this by not only focusing on improving tracking performance, but also on improving the interpretability of the tracking results. By combining neural networks and statistical physics our algorithm can predict well-calibrated error rates that indicate the chance that a tracking feature is incorrect. These error estimates can then be used to guide manual review or automatically filter out low quality tracks.

In **chapter 4,** we use semi-automated cell tracking in combination with endpoint staining to reveal basic but until now unresolved facts about intestinal differentiation. Our results can be best summarized by the 'commit-then-sort' concept: We see that all major cell types of the intestine already commit to their fate within the stem cell zone and only then move to their proper positions.

In **chapter 5,** we dive deeper into the differentiation process by looking at the process of subtype specification in tuft cells. We discover that the two tuft subtypes exist on a continuum of states from the crypt to the villus and that individual cell change their identity when

migrating along this axis. By controlling external signaling we can enrich for the villus-located subtype in organoids, enabling in vitro study of these cells.

If cell types commit early and then 'sort', by what kind of mechanisms do they actually find their proper location? In **chapter 6,** we focus on the checkerboard-pattern of secretory and non-secretory cells in the crypt. We find that the main driver of the pattern is not the cell signaling that instructs neighboring cells to adopt different cell types. Instead, the mechanical separation of already committed secretory neighbors is what keeps secretory cell spaced apart in a fast-renewing epithelium like the intestine.

A new approach to cell tracking

The biggest methodological breakthrough of this thesis work is presented in **chapter 3,** where we describe a conceptually new approach to cell tracking. Instead of focusing solely on improving tracking performance as done previously[53,87,92], we changed the focus to the interpretability of the tracking results. Our approach to compute error rates associated with each step in a cell's track is not dependent on the particulars of our algorithm, and has the potential to be broadly used in the cell tracking field and possibly in related fields like single particle tracking[287]. We argue that such indications of tracking quality will be essential for any new cell tracking algorithm, as the use of artificial intelligence makes interpretability more difficult and new microscopy techniques will create datasets so large that full manual inspection is rendered impossible[288].

Our work also reflects the maturation of cell tracking the field. Cell tracking is quickly approaching the moment where tracking quality and robustness will be good enough to facilitate adoption beyond a handful of specialized labs. This means that the focus will (and should) move from new conceptual developments in the tracking itself to ensuring robustness, interpretability of results, user-friendliness and downstream analysis tools. The error rates we provide with our predictions are an important step in that direction. As for the other aspects; our Graphical User Interface allows convenient manual curation of cell tracks even by less computationally-inclined users, mostly thanks to work from Rutger Kok[53]. Other groups have in recent years made major strides in the visualization of cell tracking data as well[289], especially for extremely large embryological datasets that are notoriously hard to visualize[290,291]. An area that has gotten comparatively less attention is the downstream analysis of tracking results. This challenge requires the integration of the tree-like data structure of the lineages and their corresponding spatial cell positions, as well as morphological and biochemical measurements. An analysis pipeline that can perform statistical analysis on such cell tracking data would be essential for widespread adoption of cell tracking methods. Indeed, it is difficult to imagine the explosive growth of the single-cell sequencing field, if it had not been paired with the development of an ecosystem of interoperable downstream analysis tools[292].

These future innovations would benefit greatly from coordination between the different leading labs in the cell tracking field. The current lack of interoperability between different cell tracking options raises the barrier of entry for experimentalists, as they cannot iterate quickly over different methods. The fact that most of the best cell tracking methods currently available use different graphical user interfaces compounds that barrier and slows down the development of new functionalities[53,87,293]. Due to differences in tracking formats, downstream analysis pipelines do not work out-of-the-box on tracking data from different sources. Fortunately, things seem to be in flux. This summer, the community has developed a common file format, and many labs are now building their GUI on top of the Napari software[293,294]. Currently, our custom GUI is probably still better suited for the manual curation of tracking data than comparable Napari-based solutions, with the most common operations only requiring a single click or key stroke, but this might change soon as more and more developers will contribute to Napari. Apart from file conventions and GUI-centered considerations, what is still lacking are community standards regarding intermediate steps in the tracking. For instance, many tracking algorithms construct a graph representation of the tracking problem before finding the optimal tracking solution, but the actual implementation of this graph object varies widely[92,293]. In practice, this means that something like our error rate computation has to be completely rewritten before it can be integrated into another algorithm. Continuing and deepening collaboration between tool-development groups thus remains necessary.

Potential improvements for our tracking algorithm

Further improvements to our tracking algorithm could takes two main routes: improving the neural network performance within our current statistical framework or further extending the statistical framework to better account for factors like segmentation mistakes. The first route, improving the neural networks, could involve increasing the amount of training data and performing hyperparameter screens to identify their best settings. More major improvements could come from replacing the convolutional neural networks by more modern architectures, like transformer-based approaches that have the potential to better integrate information over multiple timepoints[295,296]. This would differ little from the work typically done in AI research, and might therefore be best left to domain experts in the computer vision field (although it must be noted that most major cell tracking advances have come from what we can we broadly define as ‘biology’ labs[87,92,293]). The modular structure of our algorithm could in principle incorporate these new neural network architectures without much difficulty.

In the second route, we would perfect our statistical description of the tracking problem by replacing probabilities that are now set by the user with neural network predictions. This would mean estimating at each tracing step, the chances that the cell is over- or under-segmented or undergoes cell death. The easiest of these would be predicting cell death, as this could be based on cell appearance and thus would allow reusing the neural network

architecture used for the prediction of cell division. Much trickier are assigning probabilities to the correctness of a cell detections. One could plausibly train a neural network to detect over- and under-segmentations, but this model would then be specific to the mistakes made by a certain cell detection algorithm. This would in turn necessitate the creation of a new training dataset every time we train a new cell detection model, by manually correcting its mistakes.

Instead, it would be better to aim for a neural network that both identifies potential cell centers and gives us the probability that they are correct detections. Ideally, the neural network should even give the probability that two nearby cell center detections actually represent the same cell. It should then be possible to combine both these segmentation probabilities and the link probabilities in a single probabilistic graph representation. Indeed, a method called *Ultrack* is already capable of representing different segmentation and linking options in as single graph, although it cannot yet use probabilities to describe their relative likelihoods[297]. Unfortunately, a full probabilistic description of the segmentation step is not trivial. The cell detection and the probability assignment capabilities require two qualitatively very different outputs. Most cell detection algorithms rely on a pixel-wise distance-map or labelling as an output, while the probabilities we would like to predict are defined on a graph with the cell detections as nodes.

A possible approach would be to conceptualize the whole imaging volume as a graph with adjacent pixels being connected. We could then ask the neural network, in addition to the distance-map, to identify for every pair of touching pixels if they belong to the came cell (or are both background) or to different cells. After selecting potential cell centers, we could then to compute the probability of two potential cell centers representing the same cell by evaluating the pixel-pair probabilities along a path connecting these cell centers. How exactly we would combine the individual probabilities along this path to come to an overall probability is again a non-trivial problem, but this approach at least serves as a way to combine spatial and probabilistic information in a single output.

The design of such a probabilistic cell detection algorithm together with the implementation of cell death prediction would complete our probabilistic approach, which would then have a way of assigning probabilities to all relevant events for tracking. I believe that such a framework would be a natural endpoint in the conceptual development of cell tracking approaches, after which the major focus would shift to improving performance in the individual subtasks.

Other methodological advances

Chapters 4, 5 and 6 mostly deal not with methodology but with biological results. Still, it is interesting to consider their methodological aspects as well. Indeed, **chapter 4** was originally written with a much stronger emphasis on the methodology. The in-situ staining method developed by Xuan Zhang could be key in linking live cell tracking with cell type information

in many other organoid systems as well. **Chapter 5** is a nice showcase of the potential synergies between single cell sequencing and single cell tracking. In this case, the former provides high-resolution data in terms of cell identity and is able to resolve the different subtypes and transitionary states between them. Cell tracking can only study one or two subtype markers at the same time but does provide temporal and spatial information. By carefully selecting which marker to track, we were able to combine these methodologies and reach conclusions that would be impossible to draw with either technique on its own. Again, this could serve as model for mapping differentiation in other organoid systems as well. Lastly, in **chapter 6**, we use laser cell ablation to force new cellular arrangements, which is probably the closest we can get to the cutting and pasting experiments that have taught us so much about embryology. Cell ablation is not typically done in 3D tissues, but our protocol, mainly developed by my colleague Kasper Spoelstra, uses low laser powers aimed precisely at the cell membrane to target cells for extrusion. By co-opting the natural extrusion process, we can precisely target cells for ablation in a way that is minimally invasive. These new imaging-related techniques that we have developed in our lab in the past years, might prove to be as consequential as any given scientific discovery we made using them.

Cell types

Moving beyond methodological innovations, our first major biological result is probably presented in **chapter 4.** These results can best be summarized by the 'commit-then-sort' concept: All major cell types of the intestine already commit to their fate within the stem cell zone and only then move to their proper positions. Our method of cell tracking followed by antibody staining was uniquely suited for this discovery as it does not rely on prior knowledge of markers of this early commitment. By simply 'backtracking' the final fates of the cells along their lineage trees, we could establish the moment of fate commitment using only known and mature cell type markers. Indeed, simultaneous sequencing and lineage tracing work that relied on novel early markers for the Paneth and enteroendocrine fates, validated our work *in vivo* for these specific cell types, showing that they commit to their fate within the stem cell zone while still expressing the canonical stem cell marker LGR5[173]. Subsequent work that used lineage tracing and single cell sequencing in combination with modelling similarly implied early fate commitment of all major intestinal cell types[298].

A surprising feature of our data was that sister cells typically displayed the same cell type, even for comparatively rare secretory types, implying that commitment to these types occurred in their mother cell, prior to the last division. While *in vivo* single-cell sequencing data does show expression of cell cycle markers in the early progenitors of secretory cells[173], it remains to be seen if all cell types undergo a final division after commitment as in the organoid. An alternative scenario could be that cells commit to their fate immediately after their final division, retaining some cell cycle marker expression. In the absence of a marker for the actual moment of fate determination for all cell types, intravital imaging of cells

labelled with both a nuclear marker to track them and a cell fate marker seems the only option to truly resolve this question. Interestingly, a study from 1985 already concluded that Paneth cells must almost exclusively arise (and die) in pairs to account for the distribution of Paneth cell numbers per crypt, as crypts containing an even number of Paneth cells are strongly overrepresented[299]. In 2006, the same method was deployed to show that at least 15% - but potentially all - enteroendocrine cells arise in pairs, depending on assumptions about how the sister cells leave the crypt[300]. It would be interesting to revisit this *in vivo* 'cell counting' method. With our improved understanding of intestinal differentiation, we could then use earlier cell type markers that would be more effective in determining the exact symmetric division rate.

Subtypes

In **chapter 5** we zoomed in on the differentiation process by focusing on the process of subtype specification in the tuft cell lineage. In contrast to cell type commitment, we see strong external control of subtype specification mainly by BMP signaling, which increases along the crypt-villus axis and can radically shift the distribution and expression of subtype specific markers. It is interesting to speculate why the subtype determination, in contrast to type determination, does not rely on early, largely cell-autonomous, decisions. We could imagine that the more subtle differences between subtypes necessitate a less strict separation of roles and leave more room for plasticity and responsiveness to external cues. The fact that the two tuft subtypes seem to lie on a continuum in RNA expression points in that direction as well.

In studying why the tuft subtypes are specified in this particular way, performing a similar study of the enteroendocrine lineage may prove insightful, especially because the function of these subtypes is much clearer; every subtype makes its own hormone. Research suggests that the entero-endocrine lineage employs both early bifurcations in the differentiation pathway and BMP-governed trans-differentiation along the crypt-villus axis to generate this diversity[14,147]. Interestingly, enteroendocrine cells also seem to be in control of their spatial position, with older mature cells being able to stay in the crypt for extended times[15]. The enteroendocrine lineage thus exhibits a much wider repertoire of behaviors than the tuft cell lineage, which should make it easier to evaluate why certain strategies have been favored by evolution.

This research program would mimic our experiments on the tuft cell lineage: high throughput sequencing, followed by live tracking of the most informative markers in organoids. The main challenge would be that enteroendocrine cells generally stay in the crypt during organoid culture, which probably leads to only partial maturation for some subtypes. BMP-supplementation would induce further differentiation[14], like it does for the tuft cells, but would also affect the crypt-residing mature subtypes. A full understanding of enteroendocrine differentiation would probably require new culturing techniques, like

micropatterning of BMP gradients that locally induce villus or upper-crypt like domains in a way that recapitulates the *in vivo* situation[301].

Mechanical patterning

In **chapter 6** we looked not at each cell's position along the crypt-villus axis, but at their patterning with respect to their neighbors within the crypt compartment. We showed that the pattern formed by Paneth and stem cells at the crypt bottom and that of secretory and non-secretory cells higher up in crypt is not the result of lateral inhibition as is commonly thought[16,243-245]. During lateral inhibition secretory cells use cell signaling in the form of the Notch pathway to prevent the secretory differentiation in their neighbors, instead we find that the pattern is maintained mechanically not through signaling. The stem-stem cell interface is under increased active tension, which leads to a shrinkage of their interfaces with each other and an increase of their interfaces with Paneth cells, separating them to form a 'checkerboard' pattern. Furthermore, Paneth cells are stiffer and less mobile in response to nearby cell extrusions, which further stabilizes the configuration.

Strikingly, our results fit well with recent work on the pattering of the Delta-positive hair cells in the epithelium of the inner ear[278,279]. The main difference seems to be that in this system, which is not rapidly renewing like the intestine, cell signaling is dominant and mechanics is responsible for further fine-tuning, while in the intestine our results show that the mechanical system dominates. Both the increased contractility of the Notch-positive cells (stem cells in the intestine, support cells in the ear) and the higher stiffness of the Delta-positive are conserved between the systems[278]. In the inner ear immunostainings have related the increased contractility to (double) myosin phosphorylation[279]. Although we were not able to perform similar staining in the intestine, experiments using a small molecule inhibitor of myosin suggest a similar mechanism. To fully understand the role of myosin contractility in pattern formation (are there for instance cell adhesion-based mechanisms involved in patterning as well?), such inhibition experiments are ill-suited. First, myosin has an enormous number of roles in the cell, most importantly in our case in driving cell division, so long term inhibition experiments remain out of reach. Secondly, the inhibitors effect the full tissue, making it hard to distinguish the effect they directly have on single interfaces from the effect of the global remodeling of the organoid. Optogenetic control of myosin contractility, recently established in intestinal organoids, would be of great help here[286]. It would allow us to disturb contractility at the single cell or even at the single interface level, isolating the distinct contributions of the different cell-cell interfaces to the pattering.

Another open question is how differences in interface contractility are controlled. Due to its conserved nature between the intestine and inner ear, it is logical to presume a role for Notch-signaling here as this is a key shared signaling pathway. Notch signaling could affect cell mechanics by inducing transcription of cell adhesion proteins and regulators of myosin activity. This would be downstream of the role of the Notch receptor's intracellular domain

as a transcription activator when cleaved off, which regulates secretory cell fate. Alternatively, Notch might act at the membrane itself, locally controlling interface mechanics. Recent studies have hinted at such capabilities of the Notch transmembrane domain[280-282]. Unfortunately, again, the typical approach of using small molecule inhibitors is poorly suited for analyzing the role of Notch signaling. Fully disrupting Notch signaling would convert all cells to a secretory fate, eliminating any pattern for us to study[302]. Mosaic organoids, where part of the cell population is rendered insensitive to Notch disruption through constitutive expression of its intracellular domain could help here by allowing some cells to stay non-secretory[283]. More speculatively one could imagine reconstituting the system synthetically, similar to the SynNotch system[303]. In that case the goal would not be to use it to control gene expression but to control the mechanics at the interface by coupling the synthetic Notch receptor to proteins that can regulate actomyosin contractility such as rhoA or Rac1[304].

No secretory progenitor

The long-term cell tracking presented in this thesis constitutes the most detailed look into the intestinal lineage hierarchy to date. One of the major surprises resulting from this tracking is the absence of a common secretory progenitor for the Paneth, enteroendocrine and goblet cells (**chapter 3**). In **chapter 5** we show that Tuft cells do not seem to derive from such a progenitor either, contrary to what is often suggested[10]. The existence of secretory progenitors is widely assumed, although their exact definition varies considerably. For the purposes of this discussion, we will define three possible definitions for a secretory progenitor. First and most strongly, a secretory progenitor might refer to a cell that can give rise to descendants of more than one secretory type[12]. Second, it might refer to a cell that can give rise to different secretory types, but because it does not divide anymore (or divides only after final fate commitment) only gives rise to one of these[157]. Third, it might be possible that the final fate is already set in the stem cells, but that all secretory cells go through very similar progenitor states, in terms of gene expression, that may appear as a single distinct secretory progenitor state.

Our research discounts the notion of a common secretory progenitor giving rise to multi-type lineages in organoids. Furthermore, it limits the duration of any common progenitor state to less than a cell cycle; after loss of stem cell-fate and before the subsequent division after which all progeny is of the same type. It also places this secretory progenitor spatially within the stem cell zone, while it was commonly thought to live within the transit amplifying compartment[10]. Recent *in vivo* single-cell RNA-seq work restricts this even further: this data shows no evidence of a common progenitor state, instead, when enough RNA is sequenced per cell, all cell types seem to have a unique progenitor state[173]. This would then even discount the weakest definition of the secretory progenitor. Hence, these results suggest that

the major intestinal cell types are generated directly from stem cells through cell type specific progenitor states.

Everything-all-at-once

In light of these recent findings, it is interesting to consider how the notion of a secretory progenitor emerged in the first place. First, the secretory cell types all share one major characteristic; the fact that they produce Notch-ligands. Production of these ligands happens early in secretory cell differentiation and functions like a binary switch[12,157], where cells either produce the Notch-receptor (stem cells and enterocytes) or its ligand (secretory cells). Activation of Notch signaling therefore seems like an obvious first event in cell fate choice that would then give rise to a secretory and absorptive progenitor population, which would then choose its final cell type based on further external signals, like WNT and MAPK signaling[126].

Another piece of evidence favoring a secretory progenitor is the existence of transcription factors, which, when knocked-out, shift cells from one secretory fate to another. For example, knocking-out GFI1 or ZNF800 depletes the number of Paneth and Goblet cells in favor of enteroendocrine cells[305,306], while a Neurogenin3 knock-out does the opposite[307,308]. Moreover, you get less Paneth cells and more goblet cells when you increase MAPK signaling[309] or make a NFKB1-knockout[310], suggesting a common Paneth/Goblet progenitor[9,311]. On a more detailed level, however, problems arise: for instance, a GFI1-knockout has a much stronger depletion effect on Paneth cells than on Goblet cells, while it is supposed to work on their shared progenitor[13,306]. At the same time, the neurogenin3 knockout converts enteroendocrine cells to goblet cells and not to Paneth cells[307,308]. The fact that it is possible to reassign secretory fates has led the field to conceptualize these knocked-out transcription factors as controlling a series of binary choices downstream from the pan-secretory progenitor. However, the fact that it is hard to build a consistent tree of secretory fates suggests major problems with this way of thinking.

In contrast to the knock-out experiments, early lineage tracing studies performed by Chang & Bjerknes revealed that if stem cells were randomly and sparsely labelled, they generally produced clones that either contained only one specific secretory type or otherwise contained absorptive cells as well. After careful modelling, they interpreted this as evidence against a secretory progenitor[260,300]. Following their conceptual model, knock-outs of transcription factors like GFI1 and Neurogenin3 then should not affect a common secretory progenitor but inhibit the further differentiation of already committed progenitors of a specific type. These ‘blocked’ cells can then transdifferentiate into other secretory cell types. Indeed when Bjerknes & Chang knock-out GFI1 they identify transdifferentiating cells with ‘intermediate’ phenotypes resembling two secretory types[13]. Although later lineage tracing studies have sometimes reached different conclusions[12], Chang & Bjerknes results have probably remained the highest quality data of their kind, both in terms of throughput and quantitative

analysis. Our live-imaging now largely validates their model of differentiation. Instead of seeing differentiation as a tree-like series of (binary) choices, it might be better to opt for an 'everything-all-at-once'-view were multiple pathways act synergistically at the same time to create a high-dimensional 'landscape' with multiple stable states corresponding to a commitment to different types. Recent work on how the interplay between the YAP/TAZ, Notch and Wnt pathways controls the Paneth cell fate during crypt formation is a promising first step in that direction[312].

On a more critical note, this long-lasting confusion surrounding the secretory progenitor indicates that the field often engages little with existing high-quality quantitative data. This type of data, be it in the form lineage tracing or cell tracking, should serve as a background against which more detailed biomolecular data and perturbation results can be interpreted. Instead, interpretation often strongly privileges the new experimental data over established 'ground-truths'. New conclusions are also often highly qualitative in nature, representing differentiation as a series of abstract arrows, not associated with physical quantities like rates, flows, or proportions. It is then difficult to check if the proposed models fit even basic facts about intestinal homeostasis. Recent high-profile work that claims the existence of a non-LGR5 stem cell population higher up in the crypt[313,314], something almost impossible to reconcile with LGR5 lineage tracing studies, is symptomatic in both the non-quantitative nature of its claims and its lack of engagement with established 'ground-truths'. The argument here is not that researchers should avoid challenging existing frameworks, but that they should at least engage seriously with the experimental results these are based on.

Future work: Lineages during regeneration and inflammation

The early and complete determination of cell fate in 'normal' intestinal homeostasis stands in somewhat of a contrast with the large plasticity of cells during damage or inflammation. After damage we see the most spectacular instances of this, with almost all cell types having some regenerative potential. For instance, in a mouse model for inflammatory bowel disease, stem cells are lost, and adult Paneth cells can take over the stem cell roll. The Paneth cells do not fully dedifferentiate, instead becoming a kind of Paneth-stem cell hybrid, but are still are capable of generating the full epithelium[4]. Other types of inflammation often induce more subtle changes to cell differentiation but still induce pronounced phenotypes. For instance, when using interleukin 4 and 13 to induce tuft cell differentiation in **chapter 5** we also increase Paneth and especially goblet cells numbers in the organoids, something which is also seen *in vivo*[190,191,315]. Many other interleukins and other well-known pro-inflammatory signals like TNF-α and IFN-γ have been shown to have a direct effect on the organoid intestinal epithelium as well[207,316,317].

Now that we have a decent grasp of differentiation in normal homeostatic organoids, a natural next step would be to study its response to damage and/or inflammation. In regard to regeneration it would be especially interesting to study the balance between early

determinism and external control during these moments of plasticity. Research in the context of intestinal tumors suggests that cells do not fully dedifferentiate after damage and their original identity remains important even if they become capable of regenerating all other cell types, suggesting strong cell autonomous effects[318].

During inflammation it would be interesting to compare lineage trees with normal differentiation. Hyperplasia of a certain cell type could be caused both by biasing differentiation or by inducing more divisions of a progenitor. Even for well-known cases, like the interleukin-4 and interleukin-13 induced tuft and goblet cell hyperplasia, this remains unknown. In general, the regulatory architecture governing inflammation is poorly understood. We know that many external signals can produce an inflammatory response, but it is unknown if these signals have largely unique effects or downstream canalization results in them working through a small set of core 'modules'. Surprisingly, recent work indicates that two different interleukins that both induce goblet cell hyperplasia work via completely different developmental pathways[207], indicating unforeseen complexity. Intestinal organoids could serve as the ideal platform to compare the effects of different inflammatory cues and to test the potential synergistic or antagonistic effect of their combinations[317]. Because many inflammatory cues directly affect the epithelium, organoids could recapitulate key inflammatory responses even without resorting to more advanced models, like co-cultures with immune cells[319]. The automated cell tracking we have developed now makes these types of screens a possibility for the first time, although more advanced downstream statistical analysis might be required to tease out the exact lineage differences between different inflammatory regimes.

Commit-then-sort

In **chapter 4** we introduced the commit-then-sort model, which poses that cells commit early to their final fate and only then move to their proper position along the crypt-villus axis. Our observation in **chapter 6** that cell mechanics and not cell differentiation regulate secretory cell position within the crypt nicely dovetails with that model. It again shows the primacy of cell movement over cell fate changes in ensuring proper patterning.

In the case of secretory cell patterning our modelling shows that patterning through mechanical means is the only available strategy; the fast division-induced rearrangements in the intestinal crypt exclude a pattering role for the relatively slow process of cell differentiation. In contrast, there is no obvious necessity for a commit-then-sort mechanism in regulating the compartmentalization along the crypt-villus axis. Indeed, it was typically believed that differentiation of a stem cell was caused by the cell being pushed out of the stem cell zone by its dividing neighbors[24] and that cell commitment was thus downstream of cell migration.

In the face of the evidence for a commit-then-sort model, then remains the question what the advantages of this strategy are. One benefit could be that by committing early, cells have

more time in their most functional mature state before they are shed again and conversely spend less time as non-functional progenitors[298,320]. This scenario does require differentiation time to be the major bottleneck instead of time it would take for a cell to exit the crypt onto the villus, otherwise a cell might just as well wait with settling on a final fate. Within a fast-renewing system like the mouse small intestine, regenerating fully every 5 days, it seems plausible that such early commitment is a necessity. Similarly, fate commitment before the final division, giving rise to the symmetric sister pairs we observe, could give cells a head-start in the differentiation process. It is also good to note that such early commitment does not seem to preclude plasticity, as is sometimes assumed[320]. During normal homeostasis, the system behaves deterministically, but mature and progenitor intestinal cells of nearly all types do seem to be able to revert to revert to stem-like state in response to damage[12,204,264,321,322].

Alternatively, one could argue that early commitment followed by cell movement is the only possible scenario due to the mechanical constraints the system faces. The crypt bottom has a very different curvature from the transit-amplifying zone (hemi-circle versus a tube), which is important in maintaining stemness[323]. Furthermore, both stem cells and Paneth cells express different cell adhesion proteins than TA-cells, that help them localize to the crypt bottom[274,324]. If one assumes that cells can only leave the stem cell zone after changing their shape and/or their adhesion proteins, it seems necessary that the start of differentiation precedes exit from the stem cell zone. Additionally, the various cell type-specific progenitor states may also already have different mechanical requirements, necessitating early cell type commitment. The Paneth cells have to move down into the stem cell zone[325] while the enterocyte and goblet progenitors move upwards with flow[326]. The entero-endocrine progenitors on the other hand have been shown to reside in the crypt for prolonged times to mature before migrating to the villus[147]. In this framework, where cell movement is highly regulated and cell-type specific, the influence of cell position on cell fate decision has to be limited, because major changes in position are only possible after changes in the cell's mechanical which in turn are downstream of cell fate decisions.

We can thus plausibly think of early commitment as a time-saving strategy suited to the needs of the rapidly regenerating intestine or a necessary consequence of the intestinal compartmentalization in mechanically different zones. It is not directly evident what kind of experiment would distinguish between these two hypotheses or even if these are the only possible explanations for early fate commitment. It is still very much unclear under which constraints intestinal differentiation operates. It would therefore be particularly insightful to apply our methodology to other self-renewing biological systems and see which features of differentiation are conserved and which ones are particular to the (small) intestine.

Expansion to other systems

In principle, our method of cell tracking in organoids could be applied to any self-renewing system for which an *ex-vivo* model is available: stomach, lung, mammary glands, and so on.

An example of a particularly good candidate would be the intrafollicular epidermis, where differentiation is relatively well understood. The bottom layer of the epidermis is attached to a basal membrane and contains continuously dividing stem cells aptly called basal cells. These feed into non-cycling skin progenitor cells in the layers above. The importance of the mechanical properties of these cell layers for differentiation is well established and it is for instance known that reducing cell adhesion to the basal membrane is essential for migrating upwards[327]. It would be very interesting to study the lineages generated by stem cells in the basal layer and how they depend on cell mechanics. Another interesting factor is that the rate of skin renewal varies considerably across different regions[328], it is thus a natural system to study the constraints rapid self-renewal imposes on the differentiation process and the effect of the cell cycle on both cell mechanics and fate decisions.

Initial lineage tracing of basal cells suggested that actually most stem cells divide asymmetrically, with one daughter leaving the basal layer[329,330]. But as we have seen in the intestine, lineage tracing is not especially capable of dissecting the details of differentiation such as identifying a terminal division after fate commitment. Indeed, later studies based on serial intravital imaging of the same epidermal do show evidence of symmetric sister behavior[331-333]. But at the 12-to-24-hour framerate that is typical for these experiments it is difficult to resolve how precisely sister behavior is correlated and what the exact differences in cell behavior are between symmetric and asymmetric events. In contrast imaging *ex vivo* skin explants[334,335] or even skin organoids[336] would give much richer information at high temporal resolution, allowing the correlation of morphological information with lineage information. The validation of our cell tracking for low signal-to-noise data means that we can be confident that (semi-) automated tracking in a multilayered structure like the skin should be possible.

Within the skin there is no patterned secretory population that would allow for a comparison with the mechanical patterning of these cells in the intestine. Other well studied epithelia and organoid systems are available that do contain a Notch-negative and Notch-positive cell population. Among them are the lung epithelium, which has its ciliated cells interspersed by various secretory types (strikingly the Notch-dependent population in this case)[337,338], and the mammary gland epithelium with its Notch-negative subpopulation of dispersed estrogen-sensitive cells[284,339,340]. Both these epithelia self-renew, which suggest that a mechanical patterning mechanism similar to that in the intestine is present. On the other hand, these epithelia self-renew at considerably slower rates and lack specific stem-cell niches and dedicated sites for cell extrusion like the intestinal villus tip[339,341]. These features put different constraints on the system and might enable other ways of patterning, such as the targeted extrusion of neighboring secretory cells[342].

The mammary gland epithelium is an especially interesting case with regards to the relationship between Notch-signaling and cell mechanics. The epithelium has a bilayered architecture, with a Delta-positive basal layer and a Notch-positive luminal (apical) layer

(which in turn contains Notch-negative estrogen sensitive cells). These two layers emerge through cell signaling and lateral inhibition[283], but the layers can also form through self-sorting of the two cell types[284,343]. It would be very interesting to see how comparable this form of mechanical pattering is to the checkerboard pattern formation discussed in **chapter 6**. A key difference, at least, is the fact that in the bilayered case we not only need to account for interactions between cells but also between cells and the extracellular matrix, as its composition can control which cell type ends up on the basal side[343]. Patterning in the system thus emerges from an interplay between Delta-Notch cell signaling, mechanical cell interactions and cells' interactions with the extracellular matrix.

Beyond epithelia, recent studies have uncovered other cases where both cell movement and cell signaling are able to create the same pattern. The establishment of the anterior-posterior axis in most vertebrate embryo's is controlled by signaling gradients, most importantly Wnt, with cell mechanics responsible for further refining domain boundaries. In contrast, cell sorting is dominant in gastruloid systems, which reconstitute the axis *ex vivo*. Here an initially mixed cell population, in terms of Wnt-signalling, sorts itself in an anterior and posterior domain. Excitingly, in the Killifish embryo cells are initially randomly distributed as well and only later sort in their proper position, showing that this mechanism is not just a feature of *ex-vivo* systems[342,344].

The interplay between cell differentiation, cell signaling and cell mechanics is a complicated one, where cause and effect are often difficult to establish. This is true both on the biomolecular level, where mechanics can modulate cell signaling and *vice versa*, but also on a more abstract level, where it remains unclear whether features of the differentiation process (e.g. early commitment) are optimized for a certain goal or the result of constraints posed by other features (i.e. mechanical differences between domains). If we want to move beyond a mere description of mutual interactions and dependencies, the cross-species and cross-organ experiments like the ones suggested in this section will be key. It is only through comparison of fast versus slow-growing, motile versus non-motile, multilayered versus monolayered, that we can hope to glean some of its evolutionary logic. The methods for the joint study of cell differentiation and cell movement described in this thesis, such as automated cell tracking and its combination with other techniques like antibody staining, single cell sequencing and photoablation, will hopefully inspire work in that direction.

Concluding remark

Science can sometimes feel Sisyphean, not only during the daily drudge of optimizing protocols, but also when reading the literature and noticing that for all the enormous progress biology has made in the last decades, we are often still circling around the same questions. Even if they are relatively mundane ones about the patterning of Paneth cells or the best way to probabilistically describe a cell tracking problem, with new research papering over old truths while failing to provide new secure footing. My hope is that the work presented in this

thesis pushes the boulder over the hill sometimes. And if it falls short of that ambition, I hoped to have perched the boulder somewhere nice and stable, close the top, waiting for someone else to give it that final push.

Acknowledgements

Without a doubt, the most enjoyable thing about my PhD was working with the ‘organoid team’ at AMOLF. You are all very generous and kind people and of course very good in what you do. Without the work of Xuan and Rutger to build on, I cannot even imagine where this PhD would have ended up. It was great fun to collaborate with both of you, and I am very happy we kept in touch after you left AMOLF. Kasper, Pascal and Michael were hugely influential through their feedback and experimental suggestions every second Friday. I find myself actually missing these endless meetings more and more. In the last part of the PhD, I had the pleasure of working with Jasper, Susanna and Heidi. I enjoyed bringing you up to speed with the OrganoidTracker. You guys have been very patient throughout all the software bugs and my long-winded explanations. Other key members of our team were of course our technicians; Yvonne, Aarzo and Corinna. You were all so helpful in thinking along with my experiments. Lastly, I wanted to mention Lidewei, my only student and an exceptional one. Without her help the final chapter of this thesis would probably not exist.

Next, I want to thank Jeroen and Sander for their supervision. You gave me the freedom to chart my own path, but with enough guidance to ensure that projects actually got finished sometimes. Thanks as well for the career advice in the past months.

AMOLF is a great place to work, and it is all due to the great colleagues and friends I made there. My favorite part of the day was lunch and not just because of ‘kroketten’ and fried fish Friday. This is not an exhaustive list, but I think back fondly about my lunches with Ananya and Gouri at the CWI, the philosophical discussions I had with Burak in the AMOLF canteen (and in Portuguese bars) and all the talking with Chi, Chris, Stefanos, Tom, Evan, Timo, Vanda and more. Kathi, must get a special mention, because of all our bike rides back to Amsterdam West together and our ARCNL cappuccino-breaks.

I thank Daan for introducing me to the Nederlandse Boekengids and our small research projects deep into AMOLF history. Thanks as well, to the other ‘theory guys’, Vahe and

Menno for joining the book club, which must have been the intellectual highpoint of my time at AMOLF.

Other AMOLF highlights were the group-outings with the Tans Lab. These were great times with pub quizzes, cricket, cave visits and improvised hikes. Thank you for the great trust you placed upon me in picking tourist attractions. Dhawal, Jack, Bibi, Marija, Nebosja, Anna, Tom, Christopher, Luca you are all great people, positive and full of energy. With your enthusiasm it was never a chore to hear you talk about single molecule biophysics all these Monday mornings.

Finally, I want to thank my family, Michiel, Emma, Isa, and Tonny and Wati as well, and lastly of course Mala.

About the author

Max Betjes was born in 1994 in Amsterdam-West but largely grew up in the Amsterdam Eastern Harbor District. This meant that AMOLF was a mere 15-minute bike ride away. Matter of fact, he often passed the building on his way to football practice, wondering what the large green letters in front of it could stand for. Something that has remained a mystery even after 5 years of working there.

After finishing high school at the 4e Gymnasium, Max left for Delft to pursue an architecture degree. He quickly soured on the prospect of becoming an architect but found that the Delft atmosphere, best described by the Dutch word 'studentikoos', had its charm. On somewhat of a whim, he enrolled in the nanobiology bachelor (a joint degree with the Erasmus University) next to his architecture degree. This program, a combination of biology, physics and math turned out to be tailor-made for his interests and after a few months he turned his attention to it full-time. This was much to the delight of his mother, who was, fairly understandably in 2013, quite concerned about architectural job prospects.

Max first true research experience was during his bachelor at the lab of Hyun Youk, then at the TU Delft. His inspiring supervision cemented the idea that he wanted to do research. During his master nanobiology he did a very enjoyable thesis project with Miao-Ping Chien at the Erasmus Medical center. He also explored the world a bit, studying a semester at the ETH Zurich and visiting the lab of Arjun Raj at the University of Pennsylvania.

In the end Max decided that there was no place like home and accepted a PhD offer at AMOLF with Jeroen van Zon and Sander Tans. The result of which is the research presented here.

While looking for a PhD five years ago, Max found love as well. He now lives together with his girlfriend Mala Dengkeng, 15 minutes from AMOLF and 5 minutes from his childhood home. During the weekend they greatly enjoy walking around the neighborhood.

Publications

Pertaining to this thesis

Betjes MA, Kok RNU, Tans SJ, van Zon JS. Cell tracking with accurate error prediction. *Nat Methods.* 22(11):2400-2410 (2025) (**chapter 3**)

Buissant des Amorie JR, **Betjes MA**, Bernink JH, Hageman JH, Geurts VE, Begthel H, Laskaris D, Heinz MC, Jordens I, Vinck T, Houtekamer RM, Verlaan-Klink I, Brunner SR, van Rheenen J, Gloerich M, Clevers H, Tans SJ, van Zon JS, Snippert HJG. Intestinal tuft cell subtypes represent successive stages of maturation driven by crypt-villus signaling gradients. *Nat Commun.* 16(1):6765 (2025) (**chapter 5**)

Kok RNU, Spoelstra WK, **Betjes MA**, van Zon JS, Tans SJ. Label-free cell imaging and tracking in 3D organoids, *Cell Reports Physical Science*, 6(4): 102522. (2025)

Nguyen NTB, Gevers S, Kok RNU, Burgering LM, Neikes H, Akkerman N, **Betjes MA**, Ludikhuize MC, Gulersonmez C, Stigter ECA, Vercoulen Y, Drost J, Clevers H, Vermeulen M, van Zon JS, Tans SJ, Burgering BMT, Rodríguez Colman MJ. Lactate controls cancer stemness and plasticity through epigenetic regulation. *Cell Metab.* 37(4):903-919.e10. (2025)

Baglamis S, Sheraton VM, Meijer D, Qian H, Hoebe RA, Lenos KJ, **Betjes MA,** Tans S, van Zon J, Vermeulen L, Krawczyk PM. Using picoliter droplet deposition to track clonal competition in adherent and organoid cancer cell cultures. *Sci Rep.* 13(1):18832. (2023)

Zheng X, **Betjes MA**, Ender P, Goos YJ, Huelsz-Prince G, Clevers H, van Zon JS, Tans SJ. Organoid cell fate dynamics in space and time. *Sci Adv*. 9(33) (2023) (**chapter 4**)

Betjes MA*, Zheng X*, Kok RNU*, van Zon JS, Tans SJ. Cell Tracking for Organoids: Lessons From Developmental Biology. *Front Cell Dev Biol.* 9:675013 (2021) (**chapter 2**)

Previous work

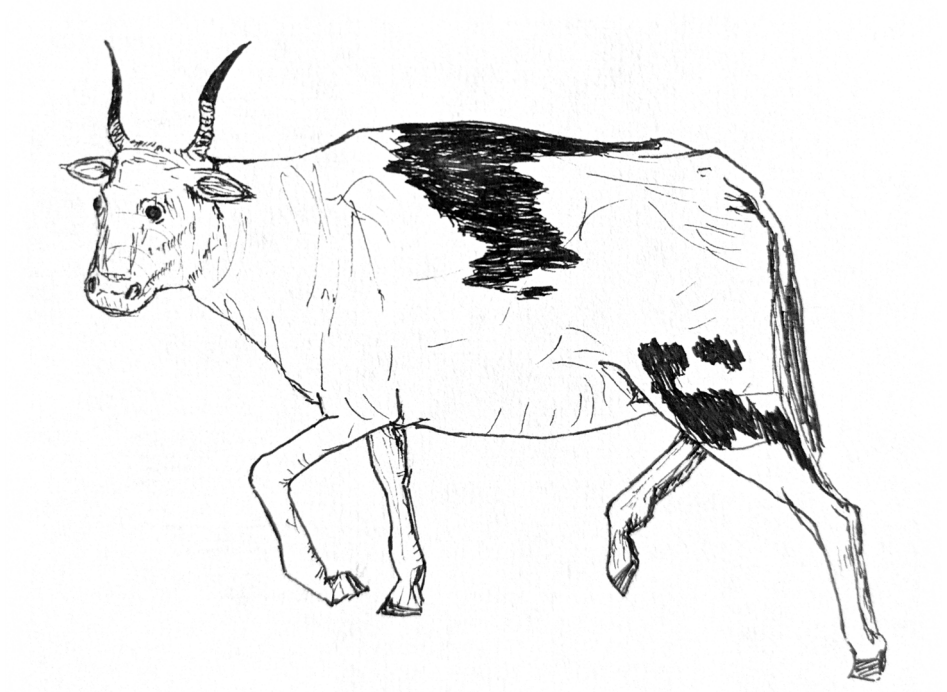
You L*, Su PR*, **Betjes M*,** Rad RG, Chou TC, Beerens C, van Oosten E, Leufkens F, Gasecka P, Muraro M, van Tol R, van Steenderen D, Farooq S, Hardillo JAU, de Jong RB, Brinks D, Chien MP. Linking the genotypes and phenotypes of cancer cells in heterogenous populations via real-time optical tagging and image analysis. *Nat Biomed Eng.* 6(5):667-675 (2022)

Maire T, Allertz T, **Betjes MA**, Youk H. Dormancy-to-death transition in yeast spores occurs due to gradual loss of gene-expressing ability. *Mol Syst Biol.* 16(11):e9245 (2020)

Miscellaneous

Max Betjes. Dino-DNA. *De Nederlandse Boekengids. (2024)*

Various letters to the ‘Volkskrant brievenrubriek’



Bibliography

1 Arias, A. M. *The Master Builder: How the New Science of the Cell is Rewriting the Story of Life*. (Basic Books, 2023).

2 Barker, N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature Reviews Molecular Cell Biology* **15**, 19-33 (2014).

3 Gersemann, M. *et al.* Differences in goblet cell differentiation between Crohn's disease and ulcerative colitis. *Differentiation; research in biological diversity* **77**, 84-94 (2009).

4 Schmitt, M. *et al.* Paneth Cells Respond to Inflammation and Contribute to Tissue Regeneration by Acquiring Stem-like Features through SCF/c-Kit Signaling. *Cell Reports* **24**, 2312-2328.e2317 (2018).

5 Ramadan, R., van Driel, M. S., Vermeulen, L. & van Neerven, S. M. Intestinal stem cell dynamics in homeostasis and cancer. *Trends in Cancer* **8**, 416-425 (2022).

6 van der Heijden, M. & Vermeulen, L. Stem cells in homeostasis and cancer of the gut. *Molecular Cancer* **18**, 66 (2019).

7 Itzkovitz, S., Blat, Irene C., Jacks, T., Clevers, H. & van Oudenaarden, A. Optimality in the Development of Intestinal Crypts. *Cell* **148**, 608-619 (2012).

8 Huelsz-Prince, G. *et al.* Mother cells control daughter cell proliferation in intestinal organoids to minimize proliferation fluctuations. *eLife* **11**, e80682 (2022).

9 Capdevila, C. *et al.* Cellular origins and lineage relationships of the intestinal epithelium. *American journal of physiology. Gastrointestinal and liver physiology* **321**, G413-g425 (2021).

10 Beumer, J. & Clevers, H. Cell fate specification and differentiation in the adult mammalian intestine. *Nature reviews. Molecular cell biology* **22**, 39-53 (2021).

11 Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262-265 (2009).

12 van Es, J. H. *et al.* Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. *Nat Cell Biol* **14**, 1099-1104 (2012).

13 Bjerknes, M. & Cheng, H. Cell Lineage metastability in Gfi1-deficient mouse intestinal epithelium. *Developmental biology* **345**, 49-63 (2010).

14 Beumer, J. *et al.* Enteroendocrine cells switch hormone expression along the crypt-to-villus BMP signalling gradient. *Nature Cell Biology* **20**, 909-916 (2018).

15 Manco, R. *et al.* Clump sequencing exposes the spatial expression programs of intestinal secretory cells. *Nat Commun* **12**, 3074 (2021).

16 Sancho, R., Cremona, C. A. & Behrens, A. Stem cell and progenitor fate in the mammalian intestine: Notch and lateral inhibition in homeostasis and disease. *EMBO Rep.* **16**, 571-581 (2015).

17 Haber, A. L. *et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333-339 (2017).

18 Elmentaite, R. *et al.* Cells of the human intestinal tract mapped across space and time. *Nature* **597**, 250-255 (2021).

19 Ancheta, S. *et al.* Challenges and Progress in RNA Velocity: Comparative Analysis Across Multiple Biological Contexts. *bioRxiv*, 2024.2006.2025.600667 (2024).

20 Zheng, S. C., Stein-O'Brien, G., Boukas, L., Goff, L. A. & Hansen, K. D. Pumping the brakes on RNA velocity by understanding and interpreting RNA velocity estimates. *Genome biology* **24**, 246 (2023).

21 Baron, C. S. & van Oudenaarden, A. Unravelling cellular relationships during development and regeneration using genetic lineage tracing. *Nature Reviews Molecular Cell Biology* **20**, 753-765 (2019).

22 Wagner, D. E. & Klein, A. M. Lineage tracing meets single-cell omics: opportunities and challenges. *Nature Reviews Genetics* **21**, 410-427 (2020).

23 McKenna, A. & Gagnon, J. A. Recording development with single cell dynamic lineage tracing. **146**, dev169730 (2019).

24 Snippert, H. J. *et al.* Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134-144 (2010).

25 Pei, W. *et al.* Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature* **548**, 456-460 (2017).

26 Mohme, M. *et al.* Optical Barcoding for Single-Clone Tracking to Study Tumor Heterogeneity. *Molecular therapy : the journal of the American Society of Gene Therapy* **25**, 621-633 (2017).

27 Weber, K. *et al.* RGB marking facilitates multicolor clonal cell tracking. *Nature medicine* **17**, 504-509 (2011).

28 Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D. & Klein, A. M. Lineage tracing on transcriptional landscapes links state to fate during differentiation. **367**, eaaw3381 (2020).

29 Mayr, U., Serra, D. & Liberali, P. Exploring single cells in space and time during tissue development, homeostasis and regeneration. *Development* **146**, dev176727 (2019).

30 Lindenburg, L. & Merkx, M. Engineering genetically encoded FRET sensors. *Sensors (Basel)* **14**, 11691-11713 (2014).

31 van de Moosdijk, A. A. A., van de Grift, Y. B. C., de Man, S. M. A., Zeeman, A. L. & van Amerongen, R. A novel Axin2 knock-in mouse model for visualization and lineage tracing of WNT/CTNNB1 responsive cells. *Genesis* **58**, e23387 (2020).

32 Artegiani, B. *et al.* Fast and efficient generation of knock-in human organoids using homology-independent CRISPR-Cas9 precision genome editing. *Nat Cell Biol* **22**, 321-331 (2020).

33 Pennisi, E. Development cell by cell. *Science* **362**, 1344-1345 (2018).

34 Ulman, V. *et al.* An objective comparison of cell-tracking algorithms. *Nature Methods* **14**, 1141-1152 (2017).

35 Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental Biology* **100**, 64-119 (1983).

36 Ritsma, L. *et al.* Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging. *Nature* **507**, 362-365 (2014).

37 Sidhaye, J. & Knoblich, J. A. Brain organoids: an ensemble of bioassays to investigate human neurodevelopment and disease. *Cell Death & Differentiation* **28**, 52-67 (2021).

38 Clevers, H. Modeling Development and Disease with Organoids. *Cell* **165**, 1586-1597 (2016).

39 Rios, A. C. & Clevers, H. Imaging organoids: a bright future ahead. *Nature Methods* **15**, 24-26 (2018).

40 Svensson, C. M., Medyukhina, A., Belyaev, I., Al-Zaben, N. & Figge, M. T. Untangling cell tracks: Quantifying cell migration by time lapse image data analysis. *Cytometry Part A* **93**, 357-370 (2018).

41 Amat, F. *et al.* Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data. *Nature Methods* **11**, 951-958 (2014).

42 Bao, Z. *et al.* Automated cell lineage tracing in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 2707-2712 (2006).

43 Schiegg, M. *et al.* Graphical model for joint segmentation and tracking of multiple dividing cells. *Bioinformatics* **31**, 948-956 (2015).

44 McDole, K. *et al.* In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level. *Cell* **175**, 859-876.e833 (2018).

45 McKinley, K. L. *et al.* Cellular aspect ratio and cell division mechanics underlie the patterning of cell progeny in diverse mammalian epithelia. *eLife* **7**, e36739 (2018).

46 Serra, D. *et al.* Self-organization and symmetry breaking in intestinal organoid development. *Nature* **569**, 66-72 (2019).

47 Held, M., Santeramo, I., Wilm, B., Murray, P. & Lévy, R. Ex vivo live cell tracking in kidney organoids using light sheet fluorescence microscopy. *PLoS ONE* **13** (2018).

48 Xing, F., Xie, Y., Su, H., Liu, F. & Yang, L. Deep Learning in Microscopy Image Analysis: A Survey. *IEEE Transactions on Neural Networks and Learning Systems* **29**, 4550-4568 (2018).

49 Meijering, E., Carpenter, A. E., Peng, H., Hamprecht, F. A. & Olivo-Marin, J. C. Imagining the future of bioimage analysis. *Nature Biotechnology* **34**, 1250-1255 (2016).

50 Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* **18**, 100-106 (2021).

51 Wolff, C. *et al.* Multi-view light-sheet imaging and tracking with the MaMuT software reveals the cell lineage of a direct developing arthropod limb. *eLife* **7**, 1-31 (2018).

52 Giladi, A. *et al.* Dissecting cellular crosstalk by sequencing physically interacting cells. *Nature Biotechnology* (2020).

53 Kok, R. N. U. *et al.* OrganoidTracker: Efficient cell tracking using machine learning and manual error correction. *PLoS One* **15**, e0240802 (2020).

54 Kiviet, D. J. *et al.* Stochasticity of metabolism and growth at the single-cell level. *Nature* **514**, 376-379 (2014).

55 Rodríguez-Colman, M. J. *et al.* Interplay between metabolic identities in the intestinal crypt supports stem cell function. *Nature* **543**, 424-427 (2017).

56 Viader-Llargués, O., Lupperger, V., Pola-Morell, L., Marr, C. & López-Schier, H. Live cell-lineage tracing and machine learning reveal patterns of organ regeneration. *eLife* **7**, e30823 (2018).

57 Alladin, A. *et al.* Tracking cells in epithelial acini by light sheet microscopy reveals proximity effects in breast cancer initiation. *eLife* **9**, e54066 (2020).

58 Bolhaqueiro, A. C. F. *et al.* Ongoing chromosomal instability and karyotype evolution in human colorectal cancer organoids. *Nature genetics* **51**, 824-834 (2019).

59 Okkelman, I. A., Foley, T., Papkovsky, D. B. & Dmitriev, R. I. Live cell imaging of mouse intestinal organoids reveals heterogeneity in their oxygenation. *Biomaterials* **146**, 86-96 (2017).

60 Doupe, D. P. & Perrimon, N. Visualizing and manipulating temporal signaling dynamics with fluorescence-based tools. *Sci Signal* **7**, re1 (2014).

61 Sonnen, K. F. *et al.* Modulation of Phase Shift between Wnt and Notch Signaling Oscillations Controls Mesoderm Segmentation. *Cell* **172**, 1079-1090.e1012 (2018).

62 Rosenbloom, A. B. *et al.* β-Catenin signaling dynamics regulate cell fate in differentiating neural stem cells. **117**, 28828-28837 (2020).

63 Delaune, Emilie A., François, P., Shih, Nathan P. & Amacher, Sharon L. Single-Cell-Resolution Imaging of the Impact of Notch Signaling and Mitosis on Segmentation Clock Dynamics. *Developmental Cell* **23**, 995-1005 (2012).

64 Massey, J. *et al.* Synergy with TGFβ ligands switches WNT pathway dynamics from transient to sustained during human pluripotent cell differentiation. **116**, 4989-4998 (2019).

65 Hormoz, S. *et al.* Inferring Cell-State Transition Dynamics from Lineage Trees and Endpoint Single-Cell Measurements. *Cell Systems* **3**, 419-433.e418 (2016).

66 Camsund, D. *et al.* Time-resolved imaging-based CRISPRi screening. *Nature Methods* **17**, 86-92 (2020).

67 Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).

68 Moffitt, J. R. *et al.* High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proceedings of the National Academy of Sciences* **113**, 11046 (2016).

69 Moffitt, J. R. *et al.* High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing. *Proceedings of the National Academy of Sciences* **113**, 14456-14461 (2016).

70 Grün, D. *et al.* Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**, 251-255 (2015).

71 Omerzu, M. *et al.* Three-dimensional analysis of single molecule FISH in human colon organoids. *Biol Open* **8**, bio042812 (2019).

72 Fannon, O. M., Bithell, A., Whalley, B. J. & Delivopoulos, E. A Fiber Alginate Co-culture Platform for the Differentiation of mESC and Modeling of the Neural Tube. *Front Neurosci* **14**, 524346-524346 (2021).

73 Lin, J.-R. *et al.* Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *eLife* **7**, e31657 (2018).

74 Rodriques, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science (New York, N.Y.)* **363**, 1463 (2019).

75 Lein, E., Borm, L. E. & Linnarsson, S. The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science* **358**, 64 (2017).

76 Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78 (2016).

77 Caicedo, J. C. *et al.* Data-analysis strategies for image-based cell profiling. *Nature methods* **14**, 849-863 (2017).

78 Czech, E., Aksoy, B. A., Aksoy, P. & Hammerbacher, J. Cytokit: a single-cell analysis toolkit for high dimensional fluorescent microscopy imaging. *BMC Bioinformatics* **20**, 448 (2019).

79 Rennerfeldt, D. A., Raminhos, J. S., Leff, S. M., Manning, P. & Van Vliet, K. J. Emergent heterogeneity in putative mesenchymal stem cell colonies: Single-cell time lapsed analysis. *PLOS ONE* **14**, e0213452 (2019).

80 Stadler, T., Skylaki, S., D. Kokkaliaris, K. & Schroeder, T. On the statistical analysis of single cell lineage trees. *Journal of Theoretical Biology* **439**, 160-165 (2018).

81 Yuan, M. *et al.* Alignment of Cell Lineage Trees Elucidates Genetic Programs for the Development and Evolution of Cell Types. *iScience* **23**, 101273 (2020).

82 Skylaki, S., Hilsenbeck, O. & Schroeder, T. Challenges in long-term imaging and quantification of single-cell dynamics. *Nature Biotechnology* **34**, 1137-1144 (2016).

83 Hicks, D. G., Speed, T. P., Yassin, M. & Russell, S. M. Maps of variability in cell lineage trees. *PLOS Computational Biology* **15**, e1006745 (2019).

84 Feigelman, J. *et al.* Analysis of Cell Lineage Trees by Exact Bayesian Inference Identifies Negative Autoregulation of Nanog in Mouse Embryonic Stem Cells. *Cell Syst* **3**, 480-490.e413 (2016).

85 Betjes, M. A., Zheng, X., Kok, R. N. U., van Zon, J. S. & Tans, S. J. Cell Tracking for Organoids: Lessons From Developmental Biology. *Frontiers in Cell and Developmental Biology* **9** (2021).

86 Beumer, J. & Clevers, H. Cell fate specification and differentiation in the adult mammalian intestine. *Nature Reviews Molecular Cell Biology* **22**, 39-53 (2021).

87 de Medeiros, G. *et al.* Multiscale light-sheet organoid imaging framework. *Nat Commun* **13**, 4864 (2022).

88 Zheng, X. *et al.* Organoid cell fate dynamics in space and time. *Science Advances* **9**, eadd6480 (2023).

89 He, Z. *et al.* Lineage recording in human cerebral organoids. *Nature Methods* **19**, 90-99 (2022).

90 Ender, P. *et al.* Spatiotemporal control of ERK pulse frequency coordinates fate decisions during mammary acinar morphogenesis. *Dev Cell* **57**, 2153-2167.e2156 (2022).

91 Sugawara, K., Çevrim, Ç. & Averof, M. Tracking cell lineages in 3D by incremental deep learning. *eLife* **11**, e69380 (2022).

92 Malin-Mayor, C. *et al.* Automated reconstruction of whole-embryo cell lineages by learning from sparse annotations. *Nature Biotechnology* **41**, 44-49 (2023).

93 Wait, E. *et al.* Visualization and correction of automated segmentation, tracking and lineaging from 5-D stem cell image sequences. *BMC Bioinformatics* **15**, 328 (2014).

94 Bragantini, J., Lange, M. & Royer, L. A. Large-Scale Multi-Hypotheses Cell Tracking Using Ultrametric Contours Maps. *ArXiv* **abs/2308.04526** (2023).

95 Mitrophanov, A. Y. & Borodovsky, M. Statistical significance in biological sequence analysis. *Briefings in Bioinformatics* **7**, 2-24 (2006).

96 Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).

97 Soneson, C. & Delorenzi, M. A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics* **14**, 91 (2013).

98 Maška, M. *et al.* The Cell Tracking Challenge: 10 years of objective benchmarking. *Nature Methods* **20**, 1010-1020 (2023).

99 Haubold, C., Aleš, J., Wolf, S. & Hamprecht, F. A. in *Computer Vision – ECCV 2016.* (eds Bastian Leibe, Jiri Matas, Nicu Sebe, & Max Welling) 566-582 (Springer International Publishing).

100 Guo, C., Pleiss, G., Sun, Y. & Weinberger, K. Q. On Calibration of Modern Neural Networks. arXiv:1706.04599 (2017). <https://ui.adsabs.harvard.edu/abs/2017arXiv170604599G>.

101 Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T. & Ronneberger, O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. arXiv:1606.06650 (2016). <https://ui.adsabs.harvard.edu/abs/2016arXiv160606650C>.

102 Platt, J. C. Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. (1999).

103 Mazzaferri, J., Roy, J., Lefrancois, S. & Costantino, S. Adaptive settings for the nearest-neighbor particle tracking algorithm. *Bioinformatics* **31**, 1279-1285 (2014).

104 Löffler, K., Scherr, T. & Mikut, R. A graph-based cell tracking algorithm with few manually tunable parameters and automated segmentation error correction. *PLOS ONE* **16**, e0249257 (2021).

105 Dietrich, F. Bayesian group belief. *Social Choice and Welfare* **35**, 595-626 (2010).

106 Dietrich, F. & List, C. in *The Oxford Handbook of Probability and Philosophy* (eds Alan Hájek & Christopher Hitchcock) 0 (Oxford University Press, 2016).

107 Dey, T. *et al.* Survival analysis—time-to-event data and censoring. *Nature Methods* **19**, 906-908 (2022).

108 Aspert, T., Hentsch, D. & Charvin, G. DetecDiv, a generalist deep-learning platform for automated cell division tracking and survival analysis. *Elife* **11** (2022).

109 Mohammadi, F. *et al.* A lineage tree-based hidden Markov model quantifies cellular heterogeneity and plasticity. *Communications Biology* **5**, 1258 (2022).

110 Nunley, H. *et al.* Nuclear instance segmentation and tracking for preimplantation mouse embryos. *Development* **151** (2024).

111 Hirsch, P. *et al.* Tracking by weakly-supervised learning and graph optimization for whole-embryo C. elegans lineages. arXiv:2208.11467 (2022). <https://ui.adsabs.harvard.edu/abs/2022arXiv220811467H>.

112 Murray, J. I. *et al.* Automated analysis of embryonic gene expression with cellular resolution in C. elegans. *Nature Methods* **5**, 703-709 (2008).

113 Matula, P. *et al.* Cell Tracking Accuracy Measurement Based on Comparison of Acyclic Oriented Graphs. *PLOS ONE* **10**, e0144959 (2015).

114 Ulicna, K., Vallardi, G., Charras, G. & Lowe, A. R. Automated Deep Lineage Tree Analysis Using a Bayesian Single Cell Tracking Approach. *Frontiers in Computer Science* **3** (2021).

115 Turetken, E., Wang, X., Becker, C., Haubold, C. & Fua, P. Network Flow Integer Programming to Track Elliptical Cells in Time-Lapse Sequences. *IEEE Transactions on Medical Imaging* **PP**, 1-1 (2016).

116 Moen, E. *et al.* Accurate cell tracking and lineage construction in live-cell imaging experiments with deep learning. (2019).

117 Wen, C. *et al.* 3DeeCellTracker, a deep learning-based pipeline for segmenting and tracking cells in 3D time lapse images. *eLife* **10**, e59187 (2021).

118 Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods* **18**, 100-106 (2021).

119 Villars, A., Letort, G., Valon, L. & Levayer, R. DeXtrusion: automatic recognition of epithelial cell extrusion through machine learning in vivo. *Development* **150** (2023).

120 Roudot, P. *et al.* u-track3D: Measuring, navigating, and validating dense particle trajectories in three dimensions. *Cell Reports Methods* **3**, 100655 (2023).

121 Arbelle, A., Reyes, J., Chen, J.-Y., Lahav, G. & Riklin Raviv, T. A probabilistic approach to joint cell tracking and segmentation in high-throughput microscopy videos. *Medical Image Analysis* **47**, 140-152 (2018).

122 Tian, C., Yang, C. & Spencer, S. EllipTrack: A Global-Local Cell-Tracking Pipeline for 2D Fluorescence Time-Lapse Microscopy. *Cell Reports* **32**, 107984 (2020).

123 Magnusson, K. E., Jalden, J., Gilbert, P. M. & Blau, H. M. Global linking of cell tracks using the Viterbi algorithm. *IEEE Trans Med Imaging* **34**, 911-929 (2015).

124 Zheng, X. *et al.* Following cell type transitions in space and time by combining live-cell tracking and endpoint cell identity in intestinal organoids. *bioRxiv*, 2022.2006.2027.497728 (2022).

125 Krotenberg Garcia, A. *et al.* Active elimination of intestinal cells drives oncogenic growth in organoids. *Cell Reports* **36**, 109307 (2021).

126 Basak, O. *et al.* Induced Quiescence of Lgr5+ Stem Cells in Intestinal Organoids Enables Differentiation of Hormone-Producing Enteroendocrine Cells. *Cell Stem Cell* **20**, 177-190.e174 (2017).

127 Verissimo, C. S. *et al.* Targeting mutant RAS in patient-derived colorectal cancer organoids by combinatorial drug screening. *eLife* **5**, e18489 (2016).

128 Barbáchano, A. *et al.* Organoids and Colorectal Cancer. *Cancers (Basel)* **13** (2021).

129 Lukonin, I., Zinner, M. & Liberali, P. Organoids in image-based phenotypic chemical screens. *Experimental & Molecular Medicine* **53**, 1495-1502 (2021).

130 Höfener, H. *et al.* Deep learning nuclei detection: A simple approach can deliver state-of-the-art results. *Computerized Medical Imaging and Graphics* **70**, 43-52 (2018).

131 Liu, R. *et al.* An Intriguing Failing of Convolutional Neural Networks and the CoordConv Solution. arXiv:1807.03247 (2018). <https://ui.adsabs.harvard.edu/abs/2018arXiv180703247L>.

132 Magnusson, K. E. G., Jaldén, J., Gilbert, P. M. & Blau, H. M. Global Linking of Cell Tracks Using the Viterbi Algorithm. *IEEE Transactions on Medical Imaging* **34**, 911-929 (2015).

133 Hinton, G. E. in *1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470).* 1-6 vol.1.

134 Narayana, M. & Haverkamp, D. in *2007 IEEE Conference on Computer Vision and Pattern Recognition.* 1-8.

135 Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262-265 (2009).

136 Lancaster, M. A. *et al.* Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373-379 (2013).

137 Miller, A. J. *et al.* Generation of lung organoids from human pluripotent stem cells in vitro. *Nature Protocols* **14**, 518-540 (2019).

138 Hu, H. *et al.* Long-Term Expansion of Functional Mouse and Human Hepatocytes as 3D Organoids. *Cell* **175**, 1591-1606.e1519 (2018).

139 Lukonin, I. *et al.* Phenotypic landscape of intestinal organoid regeneration. *Nature* **586**, 275-280 (2020).

140 Camp, J. G. *et al.* Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proceedings of the National Academy of Sciences* **112**, 15672-15677 (2015).

141 Kretzschmar, K. & Watt, Fiona M. Lineage Tracing. *Cell* **148**, 33-45 (2012).

142 Snippert, H. J. *et al.* Intestinal Crypt Homeostasis Results from Neutral Competition between Symmetrically Dividing Lgr5 Stem Cells. *Cell* **143**, 134-144 (2010).

143 Tóth, B., Ben-Moshe, S., Gavish, A., Barkai, N. & Itzkovitz, S. Early commitment and robust differentiation in colonic crypts. *Molecular Systems Biology* **13**, 902 (2017).

144 Sanman, L. E. *et al.* Transit-Amplifying Cells Coordinate Changes in Intestinal Epithelial Cell-Type Composition. *Developmental Cell* **56**, 356-365.e359 (2021).

145 Yue, Y. *et al.* Long-term, in toto live imaging of cardiomyocyte behaviour during mouse ventricle chamber formation at single-cell resolution. *Nat Cell Biol* **22**, 332-340 (2020).

146 Tallapragada, N. P. *et al.* Inflation-collapse dynamics drive patterning and morphogenesis in intestinal organoids. *Cell Stem Cell* **28**, 1516-1532.e1514 (2021).

147 Gehart, H. *et al.* Identification of Enteroendocrine Regulators by Real-Time Single-Cell Differentiation Mapping. *Cell* **176**, 1158-1173.e1116 (2019).

148 Artegiani, B. *et al.* Fast and efficient generation of knock-in human organoids using homology-independent CRISPR–Cas9 precision genome editing. *Nat Cell Biol* **22**, 321-331 (2020).

149 van der Flier, L. G. & Clevers, H. Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium. *Annual Review of Physiology* **71**, 241-260 (2009).

150 Capdevila, C. *et al.* Cellular origins and lineage relationships of the intestinal epithelium. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **321**, G413-G425 (2021).

151 Cheng, H. & Leblond, C. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine V. Unitarian theory of the origin of the four epithelial cell types. *American Journal of Anatomy* **141**, 537-561 (1974).

152 Fevr, T., Robine, S., Louvard, D. & Huelsken, J. Wnt/β-Catenin Is Essential for Intestinal Homeostasis and Maintenance of Intestinal Stem Cells. *Molecular and Cellular Biology* **27**, 7551-7559 (2007).

153 Spit, M., Koo, B. K. & Maurice, M. M. Tales from the crypt: intestinal niche signals in tissue renewal, plasticity and cancer. *Open Biol* **8** (2018).

154 Fre, S. *et al.* Notch signals control the fate of immature progenitor cells in the intestine. *Nature* **435**, 964-968 (2005).

155 Buczacki, S. J. A. *et al.* Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* **495**, 65-69 (2013).

156 Basak, O. *et al.* Mapping early fate determination in Lgr5+ crypt stem cells using a novel Ki67-RFP allele. *The EMBO Journal* **33**, 2057-2068 (2014).

157 Stamataki, D. *et al.* Delta1 Expression, Cell Cycle Exit, and Commitment to a Specific Secretory Fate Coincide within a Few Hours in the Mouse Intestinal Stem Cell System. *PLOS ONE* **6**, e24484 (2011).

158 Heuberger, J. *et al.* Shp2/MAPK signaling controls goblet/paneth cell fate decisions in the intestine. *Proceedings of the National Academy of Sciences* **111**, 3472-3477 (2014).

159 Shroyer, N. F., Wallis, D., Venken, K. J., Bellen, H. J. & Zoghbi, H. Y. Gfi1 functions downstream of Math1 to control intestinal secretory cell subtype allocation and differentiation. *Genes & development* **19**, 2412-2417 (2005).

160 van Es, J. H. *et al.* Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. *Nature cell biology* **14**, 1099-1104 (2012).

161 Schonhoff, S. E., Giel-Moloney, M. & Leiter, A. B. Neurogenin 3-expressing progenitor cells in the gastrointestinal tract differentiate into both endocrine and non-endocrine cell types. *Developmental Biology* **270**, 443-454 (2004).

162 Clevers, H. The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274-284 (2013).

163 Bonis, V., Rossell, C. & Gehart, H. The Intestinal Epithelium – Fluid Fate and Rigid Structure From Crypt Bottom to Villus Tip. *Frontiers in Cell and Developmental Biology* **9** (2021).

164 van der Flier, L. G., Haegebarth, A., Stange, D. E., van de Wetering, M. & Clevers, H. OLFM4 Is a Robust Marker for Stem Cells in Human Intestine and Marks a Subset of Colorectal Cancer Cells. *Gastroenterology* **137**, 15-17 (2009).

165 Louthan, O. Chromogranin a in physiology and oncology. *Folia Biol (Praha)* **57**, 173-181 (2011).

166 Chan, C. W. M. *et al.* Gastrointestinal differentiation marker Cytokeratin 20 is regulated by homeobox gene CDX1. *Proceedings of the National Academy of Sciences* **106**, 1936 (2009).

167 Bel, S. *et al.* Paneth cells secrete lysozyme via secretory autophagy during bacterial infection of the intestine. *Science (New York, N.Y.)* **357**, 1047-1052 (2017).

168 Kim, M., Fevre, C., Lavina, M., Disson, O. & Lecuit, M. Live Imaging Reveals Listeria Hijacking of E-Cadherin Recycling as It Crosses the Intestinal Barrier. *Current Biology* **31**, 1037-1047.e1034 (2021).

169 Sato, T. *et al.* Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469**, 415-418 (2011).

170 Yin, X. *et al.* Niche-independent high-purity cultures of Lgr5+ intestinal stem cells and their progeny. *Nature Methods* **11**, 106-112 (2014).

171 Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003-1007 (2007).

172 Yang, Q., Bermingham Na Fau - Finegold, M. J., Finegold Mj Fau - Zoghbi, H. Y. & Zoghbi, H. Y. Requirement of Math1 for secretory cell lineage commitment in the mouse intestine.

173 Böttcher, A. *et al.* Non-canonical Wnt/PCP signalling regulates intestinal stem cell lineage priming towards enteroendocrine and Paneth cell fates. *Nature Cell Biology* **23**, 23-31 (2021).

174 Batlle, E. *et al.* β-Catenin and TCF Mediate Cell Positioning in the Intestinal Epithelium by Controlling the Expression of EphB/EphrinB. *Cell* **111**, 251-263 (2002).

175 Azkanaz, M. *et al.* Retrograde movements determine effective stem cell numbers in the intestine. *Nature* **607**, 548-554 (2022).

176 Zhao, M. L. *et al.* Molecular Competition in G1 Controls When Cells Simultaneously Commit to Terminally Differentiate and Exit the Cell Cycle. *Cell Reports* **31**, 107769 (2020).

177 Sancho, R., Cremona, C. A. & Behrens, A. Stem cell and progenitor fate in the mammalian intestine: Notch and lateral inhibition in homeostasis and disease. *EMBO Rep* **16**, 571-581 (2015).

178 Farin, H. F., Van Es Jh Fau - Clevers, H. & Clevers, H. Redundant sources of Wnt regulate intestinal stem cells and promote formation of Paneth cells.

179 Kabiri Z Fau - Greicius, G. *et al.* Stroma provides an intestinal stem cell niche in the absence of epithelial Wnts.

180 Kim, T. H., Escudero S Fau - Shivdasani, R. A. & Shivdasani, R. A. Intact function of Lgr5 receptor-expressing intestinal stem cells in the absence of Paneth cells.

181 van Es, J. H. *et al.* Enteroendocrine and tuft cells support Lgr5 stem cells on Paneth cell depletion. *Proceedings of the National Academy of Sciences* **116**, 26599-26605 (2019).

182 Durand, A. *et al.* Functional intestinal stem cells after Paneth cell ablation induced by the loss of transcription factor Math1 (Atoh1).

183 Lindeboom, R. G. H. *et al.* Integrative multi-omics analysis of intestinal organoid differentiation. *Molecular Systems Biology* **14**, e8227 (2018).

184 Schapiro, D. *et al.* histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nature Methods* **14**, 873-876 (2017).

185 Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature Methods* **11**, 417-422 (2014).

186 Gómez, H. F., Dumond, M. S., Hodel, L., Vetter, R. & Iber, D. 3D cell neighbour dynamics in growing pseudostratified epithelia. *eLife* **10**, e68135 (2021).

187 Banerjee, A., McKinley, E. T., von Moltke, J., Coffey, R. J. & Lau, K. S. Interpreting heterogeneity in intestinal tuft cell structure and function. *The Journal of clinical investigation* **128**, 1711-1719 (2018).

188 Kotas, M. E. *et al.* IL-13-programmed airway tuft cells produce PGE2, which promotes CFTR-dependent mucociliary function. *JCI insight* **7** (2022).

189 O'Keefe, R. N. *et al.* A tuft cell - ILC2 signaling circuit provides therapeutic targets to inhibit gastric metaplasia and tumor development. *Nat Commun* **14**, 6872 (2023).

190 Howitt, M. R. *et al.* Tuft cells, taste-chemosensory cells, orchestrate parasite type 2 immunity in the gut. *Science (New York, N.Y.)* **351**, 1329-1333 (2016).

191 von Moltke, J., Ji, M., Liang, H. E. & Locksley, R. M. Tuft-cell-derived IL-25 regulates an intestinal ILC2-epithelial response circuit. *Nature* **529**, 221-225 (2016).

192 Gerbe, F. *et al.* Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites. *Nature* **529**, 226-230 (2016).

193 Billipp, T. E. *et al.* Tuft cell-derived acetylcholine promotes epithelial chloride secretion and intestinal helminth clearance. *Immunity* **57**, 1243-1259.e1248 (2024).

194 Xiong, Z. *et al.* Intestinal Tuft-2 cells exert antimicrobial immunity via sensing bacterial metabolite N-undecanoylglycine. *Immunity* **55**, 686-700.e687 (2022).

195 Gerbe, F. *et al.* Distinct ATOH1 and Neurog3 requirements define tuft cells as a new secretory cell type in the intestinal epithelium. *The Journal of cell biology* **192**, 767-780 (2011).

196 Westphalen, C. B. *et al.* Long-lived intestinal tuft cells serve as colon cancer-initiating cells. *The Journal of clinical investigation* **124**, 1283-1295 (2014).

197 Bjerknes, M. *et al.* Origin of the brush cell lineage in the mouse intestinal epithelium. *Developmental biology* **362**, 194-218 (2012).

198 Gracz, A. D. *et al.* Sox4 Promotes Atoh1-Independent Intestinal Secretory Differentiation Toward Tuft and Enteroendocrine Fates. *Gastroenterology* **155**, 1508-1523.e1510 (2018).

199 Herring, C. A. *et al.* Unsupervised Trajectory Analysis of Single-Cell RNA-Seq and Imaging Data Reveals Alternative Tuft Cell Origins in the Gut. *Cell Syst* **6**, 37-51.e39 (2018).

200 Nadjsombati, M. S. *et al.* Detection of Succinate by Intestinal Tuft Cells Triggers a Type 2 Innate Immune Circuit. *Immunity* **49**, 33-41.e37 (2018).

201 Ndjim, M. *et al.* Tuft cell acetylcholine is released into the gut lumen to promote anti-helminth immunity. *Immunity* **57**, 1260-1273.e1267 (2024).

202 Biton, M. *et al.* T Helper Cell Cytokines Modulate Intestinal Stem Cell Renewal and Differentiation. *Cell* **175**, 1307-1320.e1322 (2018).

203 Zwick, R. K. *et al.* Epithelial zonation along the mouse and human small intestine defines five discrete metabolic domains. *Nat Cell Biol* **26**, 250-262 (2024).

204 Huang, L. *et al.* Tuft cells act as regenerative stem cells in the human intestine. *Nature* **634**, 929-935 (2024).

205 Harnik, Y. *et al.* A spatial expression atlas of the adult human proximal small intestine. *Nature* **632**, 1101-1109 (2024).

206 Eshleman, E. M. *et al.* Microbiota-derived butyrate restricts tuft cell differentiation via histone deacetylase 3 to modulate intestinal type 2 immunity. *Immunity* **57**, 319-332.e316 (2024).

207 Lindholm, H. T. *et al.* BMP signaling in the intestinal epithelium drives a critical feedback loop to restrain IL-13-driven tuft cell hyperplasia. *Science immunology* **7**, eabl6543 (2022).

208 Bollen, Y. *et al.* Efficient and error-free fluorescent gene tagging in human organoids without double-strand DNA cleavage. *PLoS biology* **20**, e3001527 (2022).

209 Feng, X. *et al.* Tuft cell IL-17RB restrains IL-25 bioavailability and reveals context-dependent ILC2 hypoproliferation. *Nature immunology* **26**, 567-581 (2025).

210 Beumer, J. *et al.* BMP gradient along the intestinal villus axis controls zonated enterocyte and goblet cell states. *Cell Rep* **38**, 110438 (2022).

211 Silverman, J. B., Krystofiak, E. E., Caplan, L. R., Lau, K. S. & Tyska, M. J. Organization of a cytoskeletal superstructure in the apical domain of intestinal tuft cells. *The Journal of cell biology* **223** (2024).

212 Hoover, B. *et al.* The intestinal tuft cell nanostructure in 3D. *Scientific Reports* **7**, 1652 (2017).

213 Bezençon, C. *et al.* Murine intestinal cells expressing Trpm5 are mostly brush cells and express markers of neuronal and inflammatory cells. **509**, 514-525 (2008).

214 Cheng, X., Voss, U. & Ekblad, E. Tuft cells: Distribution and connections with nerves and endocrine cells in mouse intestine. *Experimental cell research* **369**, 105-111 (2018).

215 Krasteva, G. *et al.* Cholinergic chemosensory cells in the trachea regulate breathing. **108**, 9478-9483 (2011).

216 Saqui-Salces, M. *et al.* Gastric tuft cells express DCLK1 and are expanded in hyperplasia. *Histochemistry and cell biology* **136**, 191-204 (2011).

217 Ting, H. A. & von Moltke, J. The Immune Function of Tuft Cells at Gut Mucosal Surfaces and Beyond. *Journal of immunology (Baltimore, Md. : 1950)* **202**, 1321-1329 (2019).

218 van der Linden, F. H. *et al.* A turquoise fluorescence lifetime-based biosensor for quantitative imaging of intracellular calcium. *Nat Commun* **12**, 7159 (2021).

219 Pérez-González, C. *et al.* Mechanical compartmentalization of the intestinal organoid enables crypt folding and collective cell migration. *Nat Cell Biol* **23**, 745-757 (2021).

220 Thorne, C. A. *et al.* Enteroid Monolayers Reveal an Autonomous WNT and BMP Circuit Controlling Intestinal Epithelial Growth and Organization. *Dev Cell* **44**, 624-633.e624 (2018).

221 Perniss, A. *et al.* Chemosensory Cell-Derived Acetylcholine Drives Tracheal Mucociliary Clearance in Response to Virulence-Associated Formyl Peptides. *Immunity* **52**, 683-699.e611 (2020).

222 Hendel, S. K. *et al.* Tuft Cells and Their Role in Intestinal Diseases. *Frontiers in immunology* **13**, 822867 (2022).

223 Grunddal, K. V. *et al.* Adhesion receptor ADGRG2/GPR64 is in the GI-tract selectively expressed in mature intestinal tuft cells. *Molecular metabolism* **51**, 101231 (2021).

224 Snippert, H. J., Schepers, A. G., Delconte, G., Siersema, P. D. & Clevers, H. Slide preparation for single-cell–resolution imaging of fluorescent proteins in their three-dimensional near-native environment. *Nature Protocols* **6**, 1221-1228 (2011).

225 Muraro, M. J. *et al.* A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Syst* **3**, 385-394.e383 (2016).

226 van den Brink, S. C. *et al.* Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat Methods* **14**, 935-936 (2017).

227 Hashimshony, T. *et al.* CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome biology* **17**, 77 (2016).

228 Peltzer, A., Mohr, C., Stadermann, K. B., Zwick, M. & Schmid, R. nf-core/nanostring: a pipeline for reproducible NanoString nCounter analysis. *Bioinformatics* **40**, btae019 (2024).

229 Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nature Biotechnology* **42**, 293-304 (2024).

230 Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Cambridge (Mass.))* **2**, 100141 (2021).

231 Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology* **38**, 1408-1414 (2020).

232 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550 (2014).

233 Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308 (2013).

234 Fujii, M., Matano, M., Nanki, K. & Sato, T. Efficient genetic engineering of human intestinal organoids using electroporation. *Nat Protoc* **10**, 1474-1485 (2015).

235 Betjes, M. A., Kok, R. N. U., Tans, S. J. & van Zon, J. S. Cell tracking with accurate error prediction. *Nature Methods* (2025).

236 Dekkers, J. F. *et al.* High-resolution 3D imaging of fixed and cleared organoids. *Nat Protoc* **14**, 1756-1771 (2019).

237 Pérez-González, C., Ceada, G., Matejčić, M. & Trepat, X. Digesting the mechanobiology of the intestinal epithelium. *Current Opinion in Genetics & Development* **72**, 82-90 (2022).

238 Yang, Q., Bermingham, N. A., Finegold, M. J. & Zoghbi, H. Y. Requirement of Math1 for Secretory Cell Lineage Commitment in the Mouse Intestine. **294**, 2155-2158 (2001).

239 Pellegrinet, L. *et al.* Dll1- and dll4-mediated notch signaling are required for homeostasis of intestinal stem cells. *Gastroenterology* **140**, 1230-1240.e1231-1237 (2011).

240 Farin, H. F. *et al.* Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. *Nature* **530**, 340-343 (2016).

241 Pentinmikko, N. *et al.* Cellular shape reinforces niche to stem cell signaling in the small intestine. *Sci Adv* **8**, eabm1847 (2022).

242 Sjöqvist, M. & Andersson, E. R. Do as I say, Not(ch) as I do: Lateral control of cell fate. *Dev. Biol.* **447**, 58-70 (2019).

243 Chen, K.-Y. *et al.* A Notch positive feedback in the intestinal stem cell niche is essential for stem cell self-renewal. *Mol. Syst. Biol.* **13**, 927 (2017).

244 VanDussen, K. L. & Samuelson, L. C. Mouse Atonal Homolog 1 Directs Intestinal Progenitors to Secretory Cell Rather than Absorptive Cell Fate. *Dev. Biol.* **346**, 215-223 (2010).

245 May, C. L. & Kaestner, K. H. Gut Endocrine Cell Development. *Mol. Cell. Endocrinol.* **323**, 70 (2010).

246 Firmino, J., Rocancourt, D., Saadaoui, M., Moreau, C. & Gros, J. Cell Division Drives Epithelial Cell Rearrangements during Gastrulation in Chick. *Dev Cell* **36**, 249-261 (2016).

247 Ranft, J. *et al.* Fluidization of tissues by cell division and apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20863-20868 (2010).

248 Germano, D. P. J. & Osborne, J. M. A mathematical model of cell fate selection on a dynamic tissue. *J. Theor. Biol.* **514**, 110535 (2021).

249 Oguma, T. *et al.* Analyzing the effect of cell rearrangement on Delta-Notch pattern formation. *Physical Review E* **107**, 064404 (2023).

250 Sancho, R. *et al.* Fbw7 Repression by Hes5 Creates a Feedback Loop That Modulates Notch-Mediated Intestinal and Neural Stem Cell Fate Decisions. *PLoS Biol.* **11**, e1001586 (2013).

251 Mines, R. C. *et al.* Spatial Patterning from an Integrated Wnt/β-catenin and Notch/Delta Gene Circuit. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2018**, 5022-5025 (2018).

252 Buske, P. *et al.* A comprehensive model of the spatio-temporal stem cell and tissue organisation in the intestinal crypt. *PLoS Comput. Biol.* **7**, e1001045 (2011).

253 Pin, C., Watson, A. J. M. & Carding, S. R. Modelling the spatio-temporal cell dynamics reveals novel insights on cell differentiation and proliferation in the small intestinal crypt. *PLoS One* **7**, e37115 (2012).

254 Gall, L. *et al.* Homeostasis, injury, and recovery dynamics at multiple scales in a self-organizing mouse intestinal crypt. *Elife* **12** (2023).

255 Du, H., Nie, Q. & Holmes, W. R. The Interplay between Wnt Mediated Expansion and Negative Regulation of Growth Promotes Robust Intestinal Crypt Structure and Homeostasis. *PLoS Comput. Biol.* **11**, e1004285 (2015).

256 Buske, P. *et al.* On the biomechanics of stem cell niche formation in the gut--modelling growing organoids. *FEBS J.* **279**, 3475-3487 (2012).

257 Thalheim, T. *et al.* Linking stem cell function and growth pattern of intestinal organoids. *Dev. Biol.* **433**, 254-261 (2018).

258 Buissant des Amorie, J. R. *et al.* Intestinal tuft cell subtypes represent successive stages of maturation driven by crypt-villus signaling gradients. *Nat Commun* **16**, 6765 (2025).

259 Carroll, T. D. *et al.* Interkinetic nuclear migration and basal tethering facilitates post-mitotic daughter separation in intestinal organoids. *J. Cell Sci.* **130**, 3862-3877 (2017).

260 Bjerknes, M. & Cheng, H. Clonal analysis of mouse intestinal epithelial progenitors. *Gastroenterology* **116**, 7-14 (1999).

261 Tóth, B., Ben-Moshe, S., Gavish, A., Barkai, N. & Itzkovitz, S. Early commitment and robust differentiation in colonic crypts. *Mol. Syst. Biol.* **13**, 902 (2017).

262 Choi, J. *et al.* Intestinal crypts recover rapidly from focal damage with coordinated motion of stem cells that is impaired by aging. *Scientific Reports* **8**, 10989 (2018).

263 van Es, J. H. *et al.* Enteroendocrine and tuft cells support Lgr5 stem cells on Paneth cell depletion. **116**, 26599-26605 (2019).

264 Tetteh, Paul W. *et al.* Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters. *Cell Stem Cell* **18**, 203-213 (2016).

265 Yang, Q. *et al.* Cell fate coordinates mechano-osmotic forces in intestinal crypt formation. *Nat Cell Biol* **23**, 733-744 (2021).

266 Hartl, L., Huelsz-Prince, G., van Zon, J. & Tans, S. J. Apical constriction is necessary for crypt formation in small intestinal organoids. *Developmental biology* **450**, 76-81 (2019).

267 Zallen, J. A. & Blankenship, J. T. Multicellular dynamics during epithelial elongation. *Seminars in Cell & Developmental Biology* **19**, 263-270 (2008).

268 Klein, A. M. & Simons, B. D. Universal patterns of stem cell fate in cycling adult tissues. *Development* **138**, 3103-3111 (2011).

269 Bjerknes, M. & Cheng, H. The stem-cell zone of the small intestinal epithelium. I. Evidence from paneth cells in the adult mouse. *American Journal of Anatomy* **160**, 51-63 (1981).

270 Gibson, M. C., Patel, A. B., Nagpal, R. & Perrimon, N. The emergence of geometric order in proliferating metazoan epithelia. *Nature* **442**, 1038-1041 (2006).

271 Krueger, D. *et al.* Epithelial tension controls intestinal cell extrusion. *Science (New York, N.Y.)* **389**, eadr8753 (2025).

272 Banjac, I. *et al.* Fate mapping in mouse demonstrates early secretory differentiation directly from Lgr5+ intestinal stem cells. *Developmental Cell* **60**, 1281-1289.e1286 (2025).

273 Batlle, E. *et al.* Beta-catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/ephrinB. *Cell* **111**, 251-263 (2002).

274 Holmberg, J. *et al.* EphB Receptors Coordinate Migration and Proliferation in the Intestinal Stem Cell Niche. *Cell* **125**, 1151-1163 (2006).

275 Kasirer, S. & Sprinzak, D. Interplay between Notch signaling and mechanical forces during developmental patterning processes. *Current opinion in cell biology* **91**, 102444 (2024).

276 Priya, R. *et al.* Tension heterogeneity directs form and fate to pattern the myocardial wall. *Nature* **588**, 130-134 (2020).

277 Blackie, L. *et al.* A combination of Notch signaling, preferential adhesion and endocytosis induces a slow mode of cell intercalation in the Drosophila retina. *Development* **148** (2021).

278 Cohen, R. *et al.* Mechanical forces drive ordered patterning of hair cells in the mammalian inner ear. *Nat Commun* **11**, 5137 (2020).

279 Prakash, A. *et al.* Junctional force patterning drives both positional order and planar polarity in the auditory epithelia. *Nat Commun* **16**, 3927 (2025).

280 White, M. J. *et al.* Notch1 cortical signaling regulates epithelial architecture and cell–cell adhesion. *Journal of Cell Biology* **222** (2023).

281 Polacheck, W. J. *et al.* A non-canonical Notch complex regulates adherens junctions and vascular barrier function. *Nature* **552**, 258-262 (2017).

282 Koca, Y., Vuong, L. T., Singh, J., Giniger, E. & Mlodzik, M. Notch-dependent Abl signaling regulates cell motility during ommatidial rotation in Drosophila. *Cell Reports* **41** (2022).

283 Journot, R. P. *et al.* Conserved signals orchestrate self-organization and symmetry breaking of bi-layered epithelia during development and regeneration. *bioRxiv*, 2024.2007.2021.603898 (2024).

284 Lilja, A. M. *et al.* Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. *Nature Cell Biology* **20**, 677-687 (2018).

285 Fekete, D. M., Muthukumar, S. & Karagogeos, D. Hair cells and supporting cells share a common progenitor in the avian inner ear. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **18**, 7811-7821 (1998).

286 Krueger, D. *et al.* Epithelial tension controls intestinal cell extrusion. **389**, eadr8753 (2025).

287 Roudot, P. *et al.* u-track3D: Measuring, navigating, and validating dense particle trajectories in three dimensions. *Cell reports methods* **3**, 100655 (2023).

288 Wan, Y., McDole, K. & Keller, P. J. Light-Sheet Microscopy and Its Potential for Understanding Developmental Processes. *Annual review of cell and developmental biology* **35**, 655-681 (2019).

289 Shim, C. *et al.* CellTrackVis: interactive browser-based visualization for analyzing cell trajectories and lineages. *BMC Bioinformatics* **24**, 124 (2023).

290 Huijben, T. A. P. M. *et al.* inTRACKtive — A Web-Based Tool for Interactive Cell Tracking Visualization. *bioRxiv*, 2024.2010.2018.618998 (2024).

291 Schott, B. *et al.* EmbryoMiner: A new framework for interactive knowledge discovery in large-scale cell tracking data of developing embryos. *PLOS Computational Biology* **14**, e1006128 (2018).

292 Heumos, L. *et al.* Best practices for single-cell analysis across modalities. *Nature Reviews Genetics* **24**, 550-572 (2023).

293 Bragantini, J. *et al.* Ultrack: pushing the limits of cell tracking across biological scales. 2024.2009.2002.610652 (2024).

294 Ulicna, K., Vallardi, G., Charras, G. & Lowe, A. R. Automated Deep Lineage Tree Analysis Using a Bayesian Single Cell Tracking Approach. **Volume 3 - 2021** (2021).

295 Gallusser, B. & Weigert, M. J. a. e.-p. Trackastra: Transformer-based cell tracking for live-cell microscopy. arXiv:2405.15700 (2024). <https://ui.adsabs.harvard.edu/abs/2024arXiv240515700G>.

296 O'Connor, O. M. & Dunlop, M. J. Cell-TRACTR: A transformer-based model for end-to-end segmentation and tracking of cells. *bioRxiv*, 2024.2007.2011.603075 (2024).

297 Bragantini, J. *et al.* Ultrack: pushing the limits of cell tracking across biological scales. *bioRxiv* (2024).

298 Banjac, I. *et al.* Fate mapping in mouse demonstrates early secretory differentiation directly from Lgr5+ intestinal stem cells. *Dev Cell* (2025).

299 Bjerknes, M. J. B. j. Assessment of the symmetry of stem-cell mitoses. **48**, 85-91 (1985).

300 Bjerknes, M. & Cheng, H. Neurogenin 3 and the enteroendocrine cell lineage in the adult mouse small intestinal epithelium. *Developmental biology* **300**, 722-735 (2006).

301 Larrañaga, E. *et al.* Long-range organization of intestinal 2D-crypts using exogenous Wnt3a micropatterning. *Nat Commun* **16**, 382 (2025).

302 Tian, H. *et al.* Opposing Activities of Notch and Wnt Signaling Regulate Intestinal Stem Cells and Gut Homeostasis. *Cell Reports* **11**, 33-42 (2015).

303 Morsut, L. *et al.* Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. *Cell* **164**, 780-791 (2016).

304 Baba, H., Fujita, T., Mizuno, K., Tambo, M. & Toda, S. Programming Spatial Cell Sorting by Engineering Cadherin Intracellular Activity. *ACS Synthetic Biology* **13**, 1705-1715 (2024).

305 Lin, L. *et al.* Unbiased transcription factor CRISPR screen identifies ZNF800 as master repressor of enteroendocrine differentiation. **382**, 451-458 (2023).

306 Shroyer, N. F., Wallis, D., Venken, K. J., Bellen, H. J. & Zoghbi, H. Y. Gfi1 functions downstream of Math1 to control intestinal secretory cell subtype allocation and differentiation. *Genes & development* **19**, 2412-2417 (2005).

307 Jenny, M. *et al.* Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *The EMBO Journal* **21**, 6338-6347-6347 (2002).

308 Li, H. J., Ray, S. K., Kucukural, A., Gradwohl, G. & Leiter, A. B. Reduced Neurog3 Gene Dosage Shifts Enteroendocrine Progenitor Towards Goblet Cell Lineage in the Mouse Intestine. *Cellular and Molecular Gastroenterology and Hepatology* **11**, 433-448 (2021).

309 Heuberger, J. *et al.* Shp2/MAPK signaling controls goblet/paneth cell fate decisions in the intestine. **111**, 3472-3477 (2014).

310 Brischetto, C. *et al.* NF-κB determines Paneth versus goblet cell fate decision in the small intestine. *Development* **148**, dev199683 (2021).

311 Wallaeys, C., Garcia-Gonzalez, N. & Libert, C. Paneth cells as the cornerstones of intestinal and organismal health: a primer. **15**, e16427 (2023).

312 Schwayer, C. *et al.* Cell heterogeneity and fate bistability drive tissue patterning during intestinal regeneration. 2025.2001.2014.632683 (2025).

313 Capdevila, C. *et al.* Time-resolved fate mapping identifies the intestinal upper crypt zone as an origin of Lgr5+ crypt base columnar cells. *Cell* **187**, 3039-3055.e3014 (2024).

314 Malagola, E. *et al.* Isthmus progenitor cells contribute to homeostatic cellular turnover and support regeneration following intestinal injury. *Cell* **187**, 3056-3071.e3017 (2024).

315 Fung, C. *et al.* Tuft cells mediate commensal remodeling of the small intestinal antimicrobial landscape. **120**, e2216908120 (2023).

316 Biton, M. *et al.* T Helper Cell Cytokines Modulate Intestinal Stem Cell Renewal and Differentiation. *Cell* **175**, 1307-1320.e1322 (2018).

317 Pavlidis, P. *et al.* Cytokine responsive networks in human colonic epithelial organoids unveil a molecular classification of inflammatory bowel disease. *Cell Rep* **40**, 111439 (2022).

318 Verhagen, M. P. *et al.* Non-stem cell lineages as an alternative origin of intestinal tumorigenesis in the context of inflammation. *Nature genetics* **56**, 1456-1467 (2024).

319 Bar-Ephraim, Y. E., Kretzschmar, K. & Clevers, H. Organoids in immunological research. *Nature Reviews Immunology* **20**, 279-293 (2020).

320 Tóth, B., Ben-Moshe, S., Gavish, A., Barkai, N. & Itzkovitz, S. Early commitment and robust differentiation in colonic crypts. *Mol Syst Biol* **13**, 902 (2017).

321 Yan, K. S. *et al.* Intestinal Enteroendocrine Lineage Cells Possess Homeostatic and Injury-Inducible Stem Cell Activity. *Cell Stem Cell* **21**, 78-90.e76 (2017).

322 Schmitt, M. *et al.* Paneth Cells Respond to Inflammation and Contribute to Tissue Regeneration by Acquiring Stem-like Features through SCF/c-Kit Signaling. *Cell Rep* **24**, 2312-2328.e2317 (2018).

323 Pentinmikko, N. *et al.* Cellular shape reinforces niche to stem cell signaling in the small intestine. *Sci Adv* **8**, eabm1847 (2022).

324 Batlle, E. *et al.* Beta-catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/ephrinB. *Cell* **111**, 251-263 (2002).

325 Bjerknes, M. & Cheng, H. The stem-cell zone of the small intestinal epithelium. III. Evidence from columnar, enteroendocrine, and mucous cells in the adult mouse. *The American journal of anatomy* **160**, 77-91 (1981).

326 Bjerknes, M. & Cheng, H. The stem-cell zone of the small intestinal epithelium. I. Evidence from paneth cells in the adult mouse. **160**, 51-63 (1981).

327 Biggs, L. C., Kim, C. S., Miroshnikova, Y. A. & Wickström, S. A. Mechanical Forces in the Skin: Roles in Tissue Architecture, Stability, and Function. *Journal of Investigative Dermatology* **140**, 284-290 (2020).

328 Sada, A. *et al.* Defining the cellular lineage hierarchy in the interfollicular epidermis of adult skin. *Nature Cell Biology* **18**, 619-631 (2016).

329 Mascré, G. *et al.* Distinct contribution of stem and progenitor cells to epidermal maintenance. *Nature* **489**, 257-262 (2012).

330 Clayton, E. *et al.* A single type of progenitor cell maintains normal epidermis. *Nature* **446**, 185-189 (2007).

331 Rompolas, P. *et al.* Spatiotemporal coordination of stem cell commitment during epidermal homeostasis. *Science (New York, N.Y.)* **352**, 1471-1474 (2016).

332 Mesa, K. R. *et al.* Homeostatic Epidermal Stem Cell Self-Renewal Is Driven by Local Differentiation. *Cell Stem Cell* **23**, 677-686.e674 (2018).

333 Cockburn, K. *et al.* Gradual differentiation uncoupled from cell cycle exit generates heterogeneity in the epidermal stem cell layer. *Nat Cell Biol* **24**, 1692-1700 (2022).

334 Villeneuve, C. *et al.* Mechanical forces across compartments coordinate cell shape and fate transitions to generate tissue architecture. *Nature Cell Biology* **26**, 207-218 (2024).

335 Damen, M. *et al.* High proliferation and delamination during skin epidermal stratification. *Nat Commun* **12**, 3227 (2021).

336 Lee, J. *et al.* Hair-bearing human skin generated entirely from pluripotent stem cells. *Nature* **582**, 399-404 (2020).

337 Whitsett, Jeffrey A. & Kalinichenko, Vladimir V. Notch and Basal Cells Take Center Stage during Airway Epithelial Regeneration. *Cell Stem Cell* **8**, 597-598 (2011).

338 Tsao, P. N. *et al.* Notch signaling controls the balance of ciliated and secretory cell fates in developing airways. *Development* **136**, 2297-2307 (2009).

339 Van Keymeulen, A. *et al.* Lineage-Restricted Mammary Stem Cells Sustain the Development, Homeostasis, and Regeneration of the Estrogen Receptor Positive Lineage. *Cell Reports* **20**, 1525-1532 (2017).

340 Shoker, B. S., Jarvis, C., Sibson, D. R., Walker, C. & Sloane, J. P. Oestrogen receptor expression in the normal and pre-cancerous breast. **188**, 237-244 (1999).

341 Wu, M., Zhang, X., Lin, Y. & Zeng, Y. Roles of airway basal stem cells in lung homeostasis and regenerative medicine. *Respiratory Research* **23**, 122 (2022).

342 Tsai, T. Y.-C. & Pinheiro, D. Coping with uncertainty: Challenges for robust pattern formation in dynamical tissues. *Seminars in Cell & Developmental Biology* **173**, 103629 (2025).

343 Cerchiari, A. E. *et al.* A strategy for tissue self-organization that is robust to cellular heterogeneity and plasticity. **112**, 2287-2292 (2015).

344 Fulton, T., Verd, B. & Steventon, B. The unappreciated generative role of cell movements in pattern formation. **9**, 211293 (2022).