MODELING VERTEBRATE ADULT STEM CELLS AND THEIR PROGENY

Yorick Post

MODELING VERTEBRATE ADULT STEM CELLS AND THEIR PROGENY

Yorick Post

MODELING VERTEBRATE ADULT STEM CELLS AND THEIR PROGENY

MODELLEREN VAN GEWERVELDEN VOLWASSEN STAMCELLEN EN HUN NAKOMELINGEN
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 21 september 2020 des ochtends te 9.15 uur

door

Yorick Post
geboren op 17 september 1992
te Den Haag

PROMOTOR

Prof. dr. J.C. Clevers

TABLE OF CONTENTS

GENERAL INTRODUCTION AND OUTLINE OF THE THESIS

The research presented in this thesis is focused on the development and utilization of model systems for adult stem cell research. Adult stem cells are cells that have the capacity to replace lost cells through cell division. The loss of cells by natural organ turnover or damage is observed in every organ, albeit at different rates. Extreme examples include the intestinal epithelium and the cardiomyocytes of the heart, with respective cellular turnover times of several days to nearly a lifetime. While the turnover rates and regenerative capacity of a given organ can differ between animals, the loss and replacement of cells is a phenomenon shared throughout the animal kingdom. In the last decades, several experimental model systems have been developed to investigate the function of adult stem cells. Such experimental models can be used for the identification of cells that carry adult stem cell properties within an organ, and to trace and characterize their progeny. For the research presented in this thesis, we apply *in vivo* lineage tracing and *in vitro* organoid cultures to different organs and organisms.

Lineage tracing allows for the identification of all progeny coming from a single cell. With this technique a single cell is permanently marked in a way that it will transmit the same mark to its progeny. Following this marker, often a fluorescent protein, it is possible to trace the lineage of this cell and reveal the cell types, the cell numbers and the spatial location of the progeny. The earliest lineage tracing experiments were performed using different dyes to mark a cell with a color or substance. A membrane, histone or DNA dye can be used to mark a cell population and trace its progeny upon cell division. The development of genetic lineage tracing increased the specificity and stability to mark individual cells of interest. In this approach, a recombinase enzyme (Cre) is expressed from a tissue-, cell- or gene-specific promoter. The Cre recombinase recombines two target sequences, two Lox sides (LoxP). When a stop codon is flanked by two Lox sides (LoxP-STOP-LoxP) and followed by the coding sequence of a fluorescent protein, expression of the marker-Cre gene will result in the stable and permanent expression of the fluorescent protein. Cre-mediated recombination can be made inducible by fusing it to the human estrogen receptor (CreER). Recombination will only take place upon the administration of ligands of the estrogen receptor (tamoxifen), enabling specific timing of the start of lineage tracing. Genomic location of the Cre or CreER incorporation will determine the founder cell of the lineage tracing, given that gene is expressed. For adult stem cell lineage tracing studies, marker genes have been chosen that specifically mark an adult stem cell population. More recently, the use of heritable DNA barcodes and naturally occurring somatic mutations have been used to reconstruct a cellular lineage using next generation-sequencing. The lineage tracing experiments in this thesis are performed using the fluorescent protein-based genetic lineage tracing system.

Organoid technology is a three-dimensional culture method that enables the expansion of stem cells. Adult stem cell-derived organoid cultures have self-organizing properties and recapitulate the functional and structural properties of the organ of origin. Supplying a growth factor medium containing all cues a (stem) cell receives *in vivo* by mimicking the niche enables *in vitro* expansion. Organoids can be pushed into a proliferative adult stem cell state or by withdrawing or adding certain factors into a differentiated fate. Adult stem cells were long believed to be incapable of proliferation outside the body. An exception was

the long-term expansion of human epidermal cells on feeder fibroblast layers by Howard Green and colleagues. As of now, organoids have been established from various epithelial organs and can be expanded from both healthy and diseased tissue. Patient-derived organoids open up new avenues in personalized medicine to predict drug response or developing regenerative strategies. However, organoid technology is not limited to mouse and human organs but can be applied to other model organisms exploiting the same cellular and molecular mechanisms. Examples already exist for canine, feline, equine and bovine derived tissues. Organoids can also be established from pluripotent embryonic stem cells (ES) and induced pluripotent stem cells (iPS) as a model to understand the developmental biology processes of an organ. Adult stem cell-based organoids provide a genetically tractable model to understand how established tissues maintain and repair itself. Organoid technology is complementary to *in vivo* studies in the field of adult stem cell research. The organoid research described in this thesis is based on adult stem cell-derived organoids.

In the following chapters we explore adult stem cell identity, function and the role of their progeny in different tissues from the mouse, human and snake using both lineage tracing and organoid technology.

In **Chapter 1**, we first propose a generalizable definition of adult stem cell function: the ability to replace lost tissue through cell division. The first studies of adult stem cells in solid organs heavily relied on findings in the hematopoietic system containing a rare hard-wired stem cell. We describe a number of recently characterized solid tissue stem cell systems and compare it to the design of the hematopoietic system. We conclude that tissues exhibit widely divergent strategies for maintenance and repair. Adult stem cells occur as hard-wired stem cells, larger populations of undifferentiated cells competing for niche space and by committed differentiated cells that can enter the cell cycle directly given the right cues. Therefore, an adult stem cell is defined by its function to replace lost tissue.

In **Chapter 2**, we apply the most direct definition of adult stem cell function established in chapter 1 (the ability to replace lost tissue through cell division) to examine the existence of adult stem cells in the heart. The existence of stem cells in the heart has long been debated and heavily relied on the use of specific stem cell markers from other tissues. Here, we generate an unbiased map of proliferating cells and their progeny in neonatal, adult, and post damage murine hearts using marker of proliferation Ki67. While we found a robust fibroblast response in the adult heart following injury, there is no evidence for the existence of an adult stem cell population that generates *de novo* cardiomyocytes.

In **Chapter 3**, establishment of organoid cultures from the mouse and human lacrimal gland is described. Lacrimal gland organoids can be long-term expanded and recapitulate the primary tissue. The organoids contain both proliferative adult stem cells as well as differentiated cells producing antibacterial tear components. Using a functional swelling assay of human organoids,

we identify a number of compounds that phenocopy the process of tear secretion *in vivo*. In combination with a single cell atlas of the human lacrimal gland we present an experimental platform to study the (patho-)physiology of the tear gland.

In **Chapter 4**, we describe the expansion of organoid technology beyond mouse and human to reptilian tissue. Snake venom gland adult stem cells can be expanded as organoids using mammalian growth factors that mimic the *in vivo* stem cell niche. The organoids contain both proliferating progenitors and various venom-producing cells. A regional and cellular heterogeneity of venom components is maintained in culture over multiple passages. We describe a protocol for the extraction of functionally active toxins from the venom gland organoids.

Chapter 5, summarizes the findings of this thesis and explores the implications for future adult stem cell research.

DEFINING ADULT STEM CELL FUNCTION AT ITS SIMPLEST: THE ABILITY TO REPLACE LOST CELLS THROUGH MITOSIS

1

Yorick Post[1] and Hans Clevers[1,2,3,*]

[1] Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences (KNAW) and UMC Utrecht, 3584 CT Utrecht, the Netherlands.
[2] Oncode Institute, 3584 CT Utrecht, the Netherlands.
[3] The Princess Maxima Center for Pediatric Oncology, 3584 CS Utrecht, the Netherlands.
[*] Correspondence: h.clevers@hubrecht.eu

ABSTRACT

Classical studies on hematopoiesis indicate that blood cell numbers are maintained by a rare, hard-wired, transplantable stem cell (SC). Subsequent studies in other organs have implicitly assumed that all SC hierarchies follow the design of the hematopoietic system. Lineage tracing techniques have revolutionized the study of solid tissue SCs. It thus appears that key characteristics of the hematopoietic SC hierarchy (rarity of the SC, expression of specific markers, quiescence, asymmetric division, unidirectional differentiation) are not generalizable to other tissues. It is concluded that the only generalizable definition of SC function is the ability to replace lost tissue through cell division.

Keywords

14

INTRODUCTION

Sixty years of research on the blood-forming system have led to a broad consensus on the defining characteristics of hematopoietic stem cells (HSCs) and on the layout of the hematopoietic differentiation hierarchy that is maintained by these HSCs (Morrison, Uchida and Weissman, 1995; Orkin and Zon, 2008; Laurenti and Göttgens, 2018). The HSC at the apex of the blood cell hierarchy has originally been defined by its ability to be transplantable. Successive refinements have been made through detailed surface marker-based phenotyping of HSC-enriched population, always followed by functional testing in transplant assays. These classical studies have revealed that HSCs are rare and divide infrequently. 'Quiescent' HSCs can be identified through their propensity to retain DNA labels (such as BrdU): Once incorporated during S-phase, DNA labels are not diluted when the HSC spends prolonged periods in the G_0 state (Wilson *et al.*, 2008). If HSCs do divide, the outcome is asymmetric, yielding one new stem cell and one daughter cell, typically an actively dividing progenitor (Suda, Suda and Ogawa, 1984). Thus, individual HSCs carry the unique capacity to 'self-renew' and are by definition long-lived. Migration of progenitors along the hematopoietic hierarchy is unidirectional: Daughter cells become progressively lineage-restricted through consecutive fate decisions, eventually yielding mature blood cells of the various types (red cells, white cells, platelets) that are ultimately destined to die. Taken together, the identity of the multipotent HSC is hard-wired: it is the sole cell type within the hematopoietic system that has the ability to replace lost blood cells over long periods of time.

As said, these insights have emerged first and foremost from experiments that have involved advanced cell sorting combined with transplantation, approaches that have been technically challenging in other tissues. By the time researchers became interested in the study of stem cell-driven homeostatic self-renewal or tissue repair in other tissues, insights into the hematopoietic system had become textbook knowledge. And while there may not have been a rationale to do so, the defining characteristics of HSCs have served as leitmotifs in the design and interpretation of experiments addressing the identity of stem cells in other mammalian organs.

Stem cells in solid organs can function in two complementary phenomena. Most organs (like prostate, liver or lung) consist of long-lived (> 6 months) cells with relatively little physiological tissue turn-over. Upon tissue damage however, proliferation in such organs is cranked up dramatically in order to rapidly replace the lost cells. Once normal cell numbers and cell type compositions are restored, the proliferative activity returns to its low, background level. Other organs (like the epidermis, the testis, or intestinal tract epithelia) consist of cells that physiologically show high replacement rates with cellular life times of days to months. Of note, the tissues with high physiological cell turn-over are not only exposed to external insults but can all shed their discarded cells directly to the outside world. For low-turnover organs, this direct disposal route is generally not available: cell debris will have to be removed through more complex routes involving -for instance - tissue macrophages. This may be one reason why organs with no direct outlet to the outside world invest in longer-term maintenance of their cells. Below, we aim to take the reader through a number of recently described solid

tissue stem cell systems and compare their individual characteristics to the classical example of the hematopoietic system.

The risk of generalizing the use of ASC markers

Initial attempts to identify stem cells in an unexplored tissue have historically relied heavily on the use of molecular markers taken from one of the well-studied stem cell systems. This has worked well for some markers (such as CD133, CD34 and Lgr5), but has resulted in rather disastrous situations in others. As a case in point, studies aimed to define adult myocardial stem cells have relied heavily on stem cell markers known from the HSC system, including c-KIT (Beltrami *et al.*, 2003; Ellison *et al.*, 2013), SCA-1 (Uchida *et al.*, 2013), and ABCG2, a marker of so-called side population cells (Pfister *et al.*, 2008; Yellamilli and van Berlo, 2016). This has led to a flurry of papers on this subject and even to a number of clinical studies involving transplantation of the presumed cardiac stem cells. The validity in the heart of each of these markers has more recently been disputed (Jesty *et al.*, 2012; van Berlo *et al.*, 2014, 2018; Cai and Molkentin, 2017; Vicinanza *et al.*, 2018) and several retractions have resulted (Oransky and Marcus, 2018). It may be concluded that known stem cell markers may serve as starting points for stem cell research in unexplored tissues, but this should always be followed immediately by rigorous functional studies.

Replacement of lost tissue by a large population of cycling, neutrally competing "potential" stem cells

The first solid tissue for which definitive experimental evidence of stem cell function was presented is the epidermis. Over 4 decades ago, Howard Green and coworkers demonstrated that human epidermal cells could be expanded in culture for significant periods of time on mouse fibroblast feeder layers (Rheinwald and Green, 1975). This technology was rapidly converted into a treatment for patients with extended third-degree burns. Small biopsies were taken from unaffected skin areas and -over periods of weeks- expanded on mouse fibroblast feeders into large autologous keratinocyte sheets that were then applied to the affected skin surface (O'Connor *et al.*, 1981). Green's therapy has saved many lives since its introduction and remains one of the iconic demonstrations of the potential of stem cell therapy. De Luca and co-workers have built on this work to treat a large number of patients with burn lesions to their corneas with autologous, cultured cornea sheets (Rama *et al.*, 2010). In a dramatic example, they also combined Green's approach to epidermal repair with gene therapy to cure a seven-year-old child suffering from a lethal genetic blistering disease (Hirsch *et al.*, 2017).

Howard Green's culture system was adapted in several ways to quantify stem cell numbers, defined as the ability of cells from human epidermis to grow out as clones on mouse fibroblast feeder layers (Barrandon and Green, 1987; Jones and Watt, 1993). We would argue that -like the HSC transplantation assays-, interpretations of these approaches are complicated by the experimental manipulations of the cells. Moreover, they only indirectly demonstrate stem cell potential as present in the tissue of origin. Indeed, the identity of the human epidermal

stem cell has remained a matter of great debate. Based on the observation that HSCs are DNA label-retaining, the presence of similar quiescent, DNA-label retaining stem cells in mouse interfollicular epidermis has been interpreted as evidence of stemness (Braun *et al.*, 2003). Yet beyond the observation that they can proliferate when placed in culture (Morris and Potten, 1994) there is little evidence to suggest that they are the (sole) stem cells of the epidermis (see also below). Phil Jones and colleagues were the first to show that random genetic marking of the abundant proliferative cells of the basal epidermis yielded clones that change shape and size over time, but could persist for very long periods (Clayton *et al.*, 2007). This was the first evidence to doubt that rarity, asymmetric cell division or quiescence represented generalizable characteristics of adult tissue stem cells.

Similar situations were subsequently reported using a variety of lineage tracing approaches in other tissues with actively dividing cellular compartments under physiological conditions. These include the esophagus (Doupe *et al.*, 2012), intestinal crypts (Lopez-Garcia *et al.*, 2010; Snippert *et al.*, 2010), stomach glands (Leushacke *et al.*, 2013) and the testis (Klein *et al.*, 2010). All these tissues harbor large populations of frequently proliferating cells. As in the epidermis, proliferative cells in each of these tissues have the potential of generating a long-lived lineage of cellular offspring. Yet, these constantly proliferating stem cells compete 'neutrally' for niche space (Schofield, 1978), meaning that their long-term persistence within the niche is not guaranteed: the active division of one cell may lead to the displacement of its neighbor from the niche by a process that is strictly stochastic (**Figure 1A**). As a consequence, only some of these 'potential' stem cells will persist long-term in the niche and produce additional 'potential' stem cells amongst their offspring. All other 'potential' stem cells are pushed out of the niche by their proliferative neighbors, upon which they differentiate and are cleared from the system. Superficially, these latter cells resemble short-lived progenitors, as known from the hematopoietic system. Yet, the fact that they only enjoyed a short life-span is not an intrinsic characteristic of these cells: They were simply unlucky in being pushed out of their niche, before they could push out a neighbor. Over time, the neutral competition-phenomenon shapes a tissue that constantly loses stem cell clones, while the remaining clones grow ever larger.

The actively dividing "potential" stem cells in these tissues do not fit the classic characteristics of the HSC. Thus, while epidermis, esophagus, intestinal crypt, stomach gland and testis are all endowed with a very active tissue renewal function, this function is distributed over a large population of uniformly fated cells. These cells are abundant, and constantly divide in a symmetric fashion: When a potential stem cell divides, two potential stem cells result. Whether each of these two cells remains as an actual stem cell depends on the availability of niche space. Their lifespan is not predetermined; it is subject to stochastic events. Consequently, the classical stem cell characteristics of self-renewal and longevity can be attributed to a defined, large cell population in each of these tissues. Yet, one can not single out individual cells within such a population that carry hard-wired stem cell characteristics and are guaranteed to persist lifelong. And hard-wired asymmetric stem cell division does not appear to be the dominant mechanism by which stem cell numbers are kept constant. Rather, it appears that niche size

determines the number of stem cells that are retained in the niche of solid organs (**Figure 1A** and **B**).

Quiescence

Pluripotent stem cells (embryonic stem (ES) cells as well as induced pluripotent stem (iPS) cells) can be cultured in a dish over prolonged periods of time, while being kept in a state of constant proliferation. This does not appear to hamper their ability to subsequently contribute to all tissues of a healthy mouse upon injection into a blastocyst (Evans, 2001). Adult stem cells (ASC) of model organisms, such as the stem cells of the Drosophila testis and ovary are similarly known to proliferate over the entire lifespan of the fly *in vivo* (Decotto and Spradling, 2005). Along these lines, the experiments of Green and colleagues (discussed above) clearly established that mammalian epidermal stem cells can dramatically expand in culture, yet do not reach a state of 'exhaustion' or succumb to the accumulation of deleterious DNA mutations. Despite these examples, it has widely been believed in mammalian stem cell biology that it would be advantageous for ASCs to limit the numbers of consecutive cell divisions to the absolute minimum required to feed active progenitor cells into a transit-amplifying compartment. In this
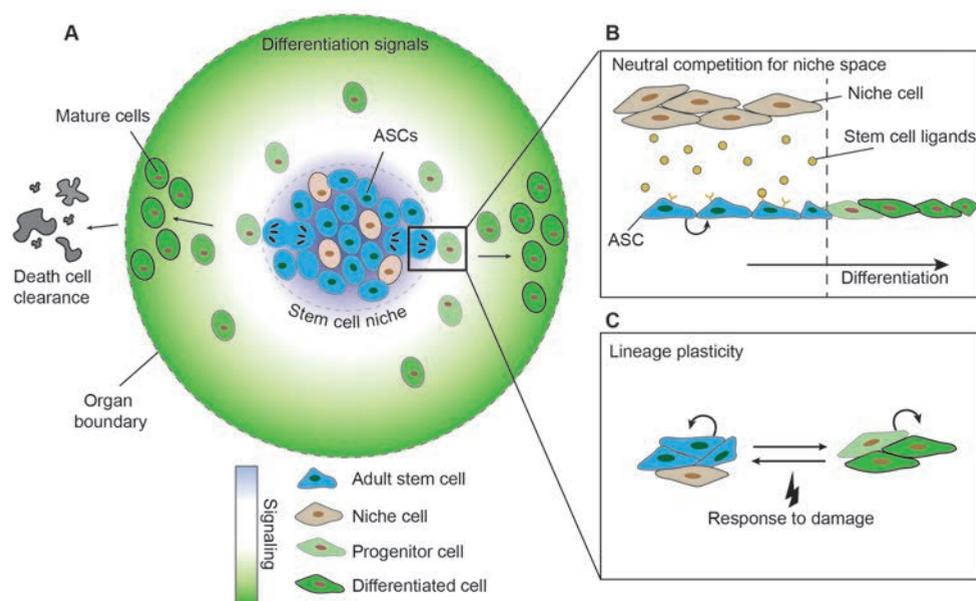
Figure 1. Niche size determines the number of stem cells. (A) Schematic depiction of Adult Stem Cells (ASCs) in blue that are restricted in cell numbers by neutral competition for niche space (indicated with dotted line). Cells that exit the stem cell niche and are beyond the reach of stem cell-promoting factors will differentiate (in green) and eventually be cleared from the organ. (B) Close-up of the niche boundary; only cells within reach of stem cell-promoting factors (in yellow) from the niche will remain as ASCs. (C) Detailed depiction of lineage plasticity. Some terminally differentiated cells have the capacity to return to a stem cell state, often induced as a response to damage.

latter scheme, the differentiated cells of a given tissue are overwhelmingly generated by transit-amplifying progenitors, rather than directly by the ASCs.

Bickenbach first reported in 1981 that 3H-thymidine, when repeatedly administered to young mice, labels DNA of slow-cycling cells in the squamous epithelia of oral mucosa and skin. These cells retained their labelled DNA into adulthood and were termed label-retaining cells (LRCs) (Bickenbach, 1981). Subsequent studies refined these methods to label epidermal LRCs, amongst others by introducing the DNA base analogue 5-bromo-2'deoxyuridine (BrdU). Although DNA label retention essentially records the proliferative history of a cell, Potten and colleagues argued on histological grounds that it specifically marks the stem cells of the epidermis. Quantification of labelled mitoses indicated the existence of two cell populations with different cell cycle times (Potten and Morris, 1988). The mouse epidermis had been noted to be 3D-organized into columns of cornified cells overlying basal cells from which they derive (Mackenzie, 1969). Slow-cycling cells were not positioned randomly in the basal layer of the epidermis: Each of these columns contained a single slowly cycling cell (Potten, 1974). Together with studies of cell cycle kinetics following radiation-induced damage repair, this led Potten and Morris to propose that the cell cycle characteristics of basal layer keratinocytes indicate their identity: Slow cycling cells are stem cells and rapidly cycling cells represent the transit-amplifying cells, the latter leaving the basal layer after a few rounds of division to terminally differentiate (Potten and Morris, 1988; Braun and Watt, 2004).

We would argue that the evidence provided for Potten's notion has been rather indirect. Moreover, the use of the LRC phenotype in a search for stem cells is complicated by another consideration. Most cells of the mammalian body do not divide, as they are terminally differentiated. Yet, they will retain DNA labels incorporated by their transit-amplifying precursors just prior to terminal differentiation. Differentiated cells will thus be scored as LRCs and may be confused for stem cells (**Figure 2**). Nevertheless, label-retaining studies have led to significant contributions to the ASC field. Cotsarelis and colleagues followed up on Bickenbach's findings and were the first to use LRC status in a hunt for new adult stem cell types. They thus identified the so called 'bulge' as the stem cell compartment of the hair follicle (Cotsarelis, Sun and Lavker, 1990). The hair follicle bulge is now one of the most studied and best-defined ASC niches, mostly through the work of Fuchs and colleagues (Hsu, Pasolli and Fuchs, 2011).

It had been postulated for some time that most HSCs are quiescent and that a succession of active HSCs guarantee life-long hematopoiesis. Weissman and colleagues were the first to give a detailed description of the cell cycle kinetics of HSCs (Cheshier *et al.*, 1999). Most murine HSCs rest in G_0 at any given time, yet all HSC regularly enter cell cycle, such that 99% of long-term (LT)-HSCs divide on average every two months. Two later studies took the analysis of the cycling behavior of HSCs a step further by monitoring the proliferative behavior of stem cells over many months *in vivo*. This revealed the presence of deeply dormant HSCs carrying the highest self-renewal potential of all blood cells (Wilson *et al.*, 2008; Foudi *et al.*, 2009).

Definitive examples of quiescent ASC types do exist in other organs also. The most iconic quiescent stem cell is probably the satellite cell of striated muscle. Healthy muscle consists of myofibers: large, contractile syncytia. After a period of postnatal proliferation of muscle tissue
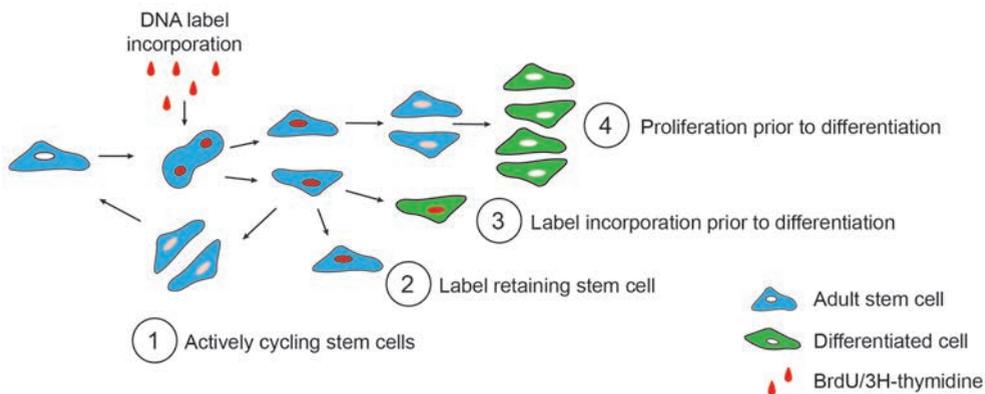
Figure 2. Interpretation of DNA label-retention. Schematic depiction of possible outcomes when using DNA labels to capture label-retaining cells (LRCs). (1) actively cycling stem cells lose their DNA label over time. (2) Quiescent stem cells retain a DNA label after its initial incorporation. (3) Immediate differentiation upon label incorporation results in label-retaining differentiated cells. (4) Multiple divisions before differentiation will dilute out a DNA label.

(White *et al.*, 2010), adult muscle stem cells are set aside to enter quiescence (Charge and Rudnicki, 2004). In response to muscle injury, these satellite cells are re-activated to proliferate and subsequently differentiate into muscle fibers to compensate for lost muscle. Some of the activated cells do not undergo myogenic differentiation, but they move back to the quiescent state in an act of self-renewal (Brack and Rando, 2012; Comai and Tajbakhsh, 2014). In one dramatic demonstration, a single satellite cell was transplanted into mouse muscle to give rise to tens of thousands of cells contributing to muscle fibers (Sacco *et al.*, 2008).

In the intestinal crypt, the proposed presence of two different stem cell types has led to significant discussion. Cheng and Leblond were the first to investigate the identity of the crypt stem cell. They noted by electron microscopy that the crypt base is not exclusively populated by Paneth cells, but that small, continuously cycling cells are wedged between these large post-mitotic cells. They termed these cells crypt base columnar (CBC) cells and -based on microscopic anatomy- proposed these as crypt stem cells (Cheng and Leblond, 1974b). They noted that injection of 3H-thymidine killed some CBC cells, that were then phagocytosed by surviving CBC cells. "Hot" phagosomes were initially only observed within surviving CBC cells but at later time points also appeared within more differentiated cells of various kinds. This experiment is probably the first -if rudimentary- lineage-tracing experiment. It supported the notion that all four main differentiated lineages derive from CBC cells (Cheng and Leblond, 1974a, 1974b)

In 1988, Winton and Ponder extended these observations by clonal labeling of crypt cells through chemical mutagenesis (Winton, Blount and Ponder, 1988). This visualized the flow of cells from crypt bottoms to villus tips as "ribbons". Using a similar chemical mutagenesis approach, Bjerknes and Cheng then showed that randomly marked crypt cells yielded long-lived as well as short-lived clones, consistent with the existence of stem cells and transit-amplifying

cells (Bjerknes and Cheng, 1999). The long-lived clones comprised all major cell lineages and presented as ribbons, always including at least one marked CBC cell, lending further support to the latter being the crypt stem cell. In 2007, we reported Lgr5 as a marker for the CBC cells and essentially confirmed all of the original notions by Leblond, Cheng and Bjerknes. By genetic lineage tracing, we demonstrated that the constantly cycling Lgr5+ stem cells are multipotent and persist life-long (Barker *et al.*, 2007). As the Lgr5+ CBC cells undergo up to 1000 consecutive cell divisions in the mouse, there appears not to exist an absolute obligation for stem cells to cycle rarely in order to preserve their proliferative potential and to minimize replicative DNA errors. In a later study, whole genome sequence was performed on clonally expanded cells from human liver, small and large intestine. The study documented that the rate of somatic mutation of single intestinal stem cells (~40 single base changes per year of life per cell) is the same as that of quiescent liver cells in man (Blokzijl *et al.*, 2016). Of note, Campbell and colleagues recently subjected single human HSCs to a similar approach and revealed a mutation rate that was in the same range (~20 single base changes per year of life per cell) (Lee-Six *et al.,* 2018).

Independent of this chain of findings, Chris Potten and colleagues reported the existence of rare LRCs located directly above the Paneth cells, at "position 4" (Potten *et al.*, 1978; Potten, Owen and Booth, 2002). These cells have since been termed "+4 cells". In these original studies, the +4 LRCs are reported to cycle every day. Thus, Chris Potten did not interpret the LRC trait of position 4 cells to reflect cellular quiescence. Rather, he proposed that the observed DNA-label retention resulted from the asymmetric segregation of old (labeled) from new (unlabeled) chromatin into rapidly dividing stem cells and their transit-amplifying daughters respectively. The "immortal strand" hypothesis had been postulated several years earlier as a mechanism to protect the stem cell genome from mutation (Cairns, 1975). It should be noted that the 'immortal strand'-mechanism has not been independently confirmed to occur in crypts since. Rather, cycling stem cells at the crypt base appear to segregate their chromosomes randomly (Escobar *et al.*, 2011; Schepers *et al.*, 2011; Steinhauser *et al.*, 2012) .

From 2008 onwards, multiple markers were proposed for Potten's +4 cells, including Bmi1 (Sangiorgi and Capecchi, 2008), Lrig1 (Powell *et al.*, 2012), Tert (Breault *et al.*, 2008) and HopX (Takeda *et al.*, 2011). All these studies erroneously cite Potten's work as describing that +4 cells are quiescent. From these studies, it was proposed that quiescent +4 cells would occupy the apex of the crypt stem cell hierarchy to give rise to the proliferative Lgr5 cells.

Winton and colleagues reassessed the nature of crypt LRCs by fluorescent marking of crypt LRCs *in vivo*. For this, they briefly expressed the stable fluorescent chromatin marker histone-2B (H2B-YFP) throughout the crypt (Buczacki *et al.*, 2013). In addition to the expected chromatin label retention by the long-lived Paneth cells, some non-Paneth-cells also retained the label for 2–3 weeks after the pulse. This second LRC type represents a nondividing Paneth/enteroendocrine precursor that apparently persists for some weeks before its terminal differentiation. To characterize these cells further, the authors devised an ingenious strategy that directly exploits the quiescent state to genetically mark the Lgr5+ LRCs. In control mice, the genetically marked LRCs disappeared over time, presumably due to their terminal differentiation. When crypts were damaged however, the LRCs generated tell-tale stem cell ribbons. Detailed molecular analyses

on sorted crypt cell populations recently confirmed Winton's notions (Jadhav *et al.*, 2017). Thus, Winton's study unified the Leblond- and Potten-models. The non-dividing +4 cells represent lineage-restricted daughters of Lgr5 cells and should not be considered stem cell *in sensu stricto* (Buczacki *et al.*, 2013). Yet, loss of the Lgr5 stem cells instructs these cells to settle at the base of the crypts and reacquire a multipotent Lgr5 stem cell phenotype (Tian *et al.,* 2011). The notion that short-lived progenitors can replace lost Lgr5 stem cells was confirmed for the secretory lineage (van Es *et al.*, 2012; Jadhav *et al.*, 2017) as well as for the more abundant, proliferative enterocyte lineage (Tetteh *et al.*, 2016). Sato and colleagues showed the same phenomenon using human colon organoids (that were CRISPR-modified to allow lineage tracing and Lgr5 stem cell removal), transplanted to mice (Shimokawa *et al.*, 2017). Importantly, these quiescent +4 cells are short-lived at the single cell level, as they will progress to full differentiation and die. Yet as a constantly replenishing compartment, +4 cells functionally build a reserve stem cell pool.

Replacement of lost tissue through proliferation of committed progenitors, or even of differentiated cells

Physiologically, cells that progress along a stem cell hierarchy become increasingly lineage-restricted, to eventually reach a terminally differentiated state, after which cell death is inevitable. The crypt example above illustrates that this process is not necessarily uni-directional. Cells underway to terminal differentiation can reacquire a multipotent stem cell state, under situations of emergency. In crypts, the cells that can do so are not yet terminally differentiated, even though they already passed through one or more lineage decision-bottle necks. More extreme cases of such plasticity have been revealed for post-mitotic cell types that are not just fate-restricted, but actually have undergone terminal differentiation (**Figure 1C**). Several recent examples are given below.

Liver

Only two epithelial cell types in the liver are endoderm-derived: the hepatocytes and the cholangiocytes. All other cell types (stellate cells, Kupffer cells, endothelium etc) derive from non-endodermal sources and immigrate into the liver after organogenesis. The liver displays a remarkable regenerative capacity in terms of recovery of mass and function upon surgical removal of a large part of the organ (partial hepatectomy), or after liver-wide chemical/infectious injury (Stanger, 2015). The response to partial hepatectomy is particularly impressive as it involves massive cell cycle entry of the remaining mature hepatocytes without apparent de-differentiation into a progenitor/stem cell-like state. The liver will thus grow back to its original size within 2 weeks after damage (Michalopoulos, 2010), while this surgical procedure can be repeated multiple times (Miyajima, Tanaka and Itoh, 2014). Of note, a low level of hepatocyte-driven homeostatic tissue renewal occurs under steady-state conditions *in vivo*: An Axin-2 lineage tracing allele has demonstrated that mature hepatocytes located around the pericentral vein (PC) drive hepatocyte renewal (Wang *et al.*, 2015). Similar genetic lineage tracing approaches have shown that periportal hepatocytes can rescue liver mass upon damage (Font-

Burgada *et al.*, 2015), as well as a subset of hepatocytes expressing high levels of telomerase (Lin *et al.*, 2018).

An alternative repair mechanism plays out when noxious agents such as toxins, viruses or hereditary defects chronically damage all hepatocytes. Under such conditions, small, proliferative cells appear near the bile duct tree. These cells typically carry early cholangiocyte markers and have been termed "oval cells". Oval cells (or atypical ductal cells expressing Epcam or Sox9) have long been believed to represent activated liver stem cells, capable of producing hepatocytes as well as cholangiocytes (Evarts *et al.*, 1987; Furuyama *et al.*, 2011; Dollé *et al.*, 2015). The hepatocyte-driven regenerative response dominates most types of liver repair (Grompe, 2014), while the *in vivo* relevance of the oval cell response as well as its cell-of-origin have long remained controversial. Indeed, several lineage tracing approaches did not support the existence of a liver stem cell with cholangiocyte characteristics (Schaub *et al.*, 2014; Yanger *et al.*, 2014). Forbes and colleagues very recently provided definitive proof for an *in vivo* oval cell response, in a mouse model in which hepatocyte proliferation was genetically blocked (Raven *et al.*, 2017). By lineage tracing, the authors demonstrated conclusively that the cell-of-origin was a cholangiocyte, rather than a professional stem cell. In sum, significant experimental efforts to identify 'professional' stem cells in the liver have not born fruit. It is most likely that the healthy liver does not contain professional stem cells and that it relies entirely on the proliferative capacity of two differentiated cell types, the hepatocyte and the cholangiocyte (**Figure 3A**). The phenomena of proliferation and lineage plasticity in liver epithelium can be recapitulated *in vitro*. Clonal organoids from a single hepatocyte (Hu *et al.,* 2018) or a single cholangiocyte (Huch *et al.,* 2015) can transdifferentiate into the other lineage, even after long-term expansion (**Figure 3B**).

Stomach epithelium

Like the intestinal lining, the stomach epithelium is organized into crypt-like structures called glands, each harboring a complete, independent stem cell hierarchy. Under physiological conditions, cell production takes places in the isthmus region, located halfway along the gland axis (Goldenring and Mills, 2017). Chief cells are long-lived, terminally differentiated cells that serve to secrete pepsinogen for protein digestion in the lumen of the stomach. They are located at the base of the epithelial glands of the stomach corpus. Chief cells contain an elaborate ER/Golgi apparatus to support their enzyme-secretory role. Two independent studies have demonstrated that Chief cells express the stem cell markers Troy (Stange *et al.*, 2013) and Lgr5 (Leushacke *et al.*, 2017), presumably because these cells sit directly on top of a Wnt source. Genetic lineage tracing using CreER alleles of these two loci revealed that, under physiological conditions, Chief cells contribute to the renewal of all other cell types of the stomach glands. However, the replacement of all cells in a gland starting from a single marked Chief cell takes in the order of a year. Replacement rates strongly accelerate upon loss of the pool of proliferative cells in the isthmus, for instance after radiation. Thus, Chief cells resemble hepatocytes and cholangiocytes in that they perform a dedicated, highly specialized function in their resident tissue, yet can 'moonlight' as stem cells, if so required (**Figure 3C**).
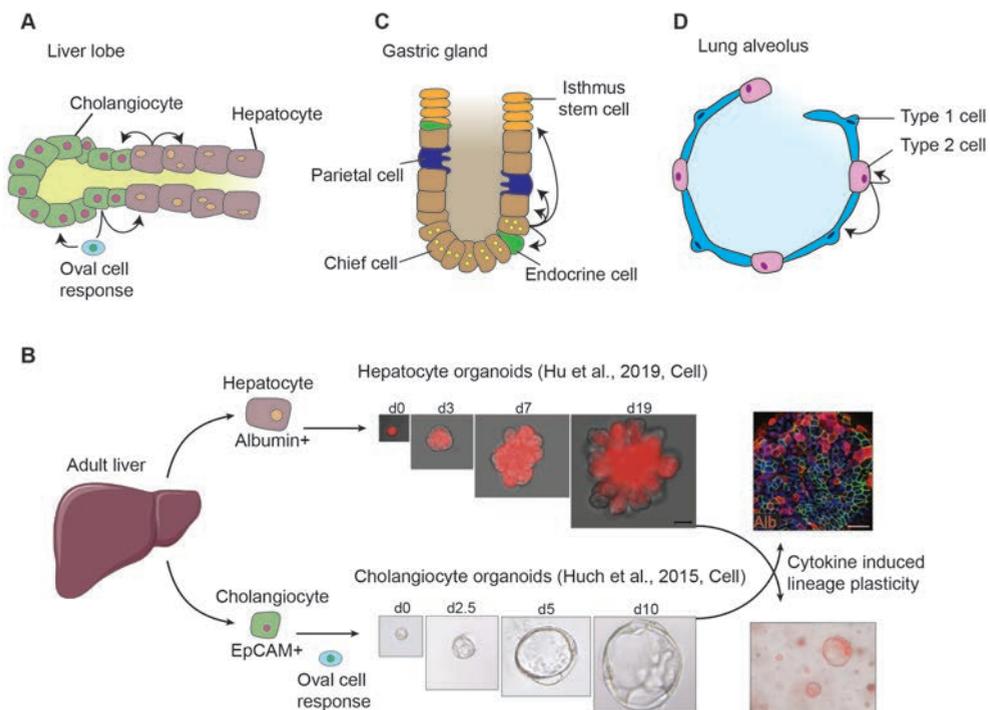
Figure 3. Differentiated cells with the capacity to proliferate and replace lost tissue. (A) The two endoderm-derived differentiated epithelial cell types of the liver (cholangiocytes and hepatocytes) both have the capacity to proliferate and regenerate both cell types. (B) Clonal organoids derived from hepatocytes or cholangiocytes can be expanded long-term and transdifferentiate into the other lineage. (C) Differentiated Chief cells are able to contribute to the renewal of all other cell types in the stomach gland. (D) In the lung alveolus, Type 2 cells can act as a reserve stem cell and generate both Type 2 and Type 1 cells.

Lung

(reviewed in Tata and Rajagopal, 2017) The adult mammalian lung can be divided into two major compartments: the airways (trachea, bronchi, bronchiole) that mediate air transport and the alveoli where exchange of O_2 and CO_2 occurs. Ciliated cells propel a thin layer of mucus towards the oral cavity, carrying with it fine-dust particles, bacteria and other debris. Secretory cells (also known as club cells or Clara cells) produce mucins as well as antimicrobial peptides, while rare neuroendocrine cells secrete various hormones controlling air- and blood-flow through the lungs. Basal cells (with an undifferentiated appearance) reside directly underneath these more linage-restricted cells and have long been believed to be the sole stem cells of the epithelium. It is now known that -under steady state conditions- fully differentiated club cells located in the basal cell-poor small airways of mice can interconvert into ciliated cells. In the large airways, ciliated cells can be formed directly by basal cells (Pardo-Saganta *et al.*, 2015). Yet, most ciliated cells in the large airway derive from club cells, thus copying the situation in the small airways (Rawlins *et al.*, 2009). Basal cells act as potent epithelial stem cells to

replenish all types of secretory and ciliated cells upon tissue damage, but fully mature club cells de-differentiate into *bona fide* basal cells upon diphtheria toxin-induced stem cell ablation (Tata *et al.*, 2013).

The other compartment of the lung, the alveolar epithelium consists of only two main cell types: The extremely flat Type 1 cells serve in gas exchange between the air in the lung and the capillaries surrounding the alveoli, while Type 2 cells produce abundant amounts of surfactant to prevent alveolar collapse by building up surface tension. In the adult mouse lung, genetic lineage-tracing has demonstrated that slow-rate proliferation of Type 2 cells generates Type 1 as well as Type 2 cells under homeostatic conditions, as well as upon damage (Barkauskas *et al.*, 2013; Desai, Brownfield and Krasnow, 2014). This activity is controlled by Notch (Finn *et al.*, 2019) and BMP signals (Chung *et al.*, 2018). Thus, Type 2 cells represent another example of a differentiated cells, 'moonlighting' as reserve stem cells (**Figure 3D**).

Should the cells in these liver-, stomach- and lung-examples, be considered as 'quiescent'? Indeed, they are predicted to retain incorporated DNA labels in their unperturbed state and as such would qualify as LRCs. Yet, these cells have not ceased to divide because they are multipotent stem cells; they have done so in the process of terminal differentiation.

Is the HCS as hard-wired as believed?

Against the backdrop of recent findings on ASC biology in solid tissues, it may be of interest to re-assess the HSC with the newly developed single cell technologies. Indeed, mapping of HSCs at single cell resolution has revealed that these cells are much more heterogeneous than assumed. Differences in marker expression, cell cycle status, lineage fate and functional outcome of transplantation is adding complexity to the apex of the lineage tree (reviewed in Crisan and Dzierzak, 2016). Camargo and colleagues recently used transposon tagging to trace single progenitor cells and decipher their fate *in vivo*. This tracing method revealed rather unexpectedly that the megakaryocyte lineage arises independently from the other blood cell types. Cells traditionally known as long-term HSCs turn out to be a significant source of megakaryocyte-restricted progenitors, suggesting this to be their predominant fate. Combined with single cell RNA sequencing, uni- and oligopotent cells were found to exist within the multipotent progenitor pool (Rodriguez-Fraticelli *et al.*, 2018). A different approach involving tracking of mature blood cells derived from a single transplanted HSC revealed the existence of a similar lineage-restricted fate of long-term HSCs towards megakaryocytes (Carrelha *et al.*, 2018).

Mapping the accumulation of somatic mutations throughout life in cells allows the reconstruction of cellular population dynamics in humans. Campbell and colleagues were the first to use this molecular mark to study relationships between HSCs and their ability to generate blood cells. Reconstruction of a phylogenetic tree based on somatic mutations confirmed the existence of adult human HSCs contributing to different lineages of the hematopoietic system. However, granulocytes and T lymphocytes did not share as many somatic mutations, again indicating an early separation of blood lineages (Lee-Six *et al.*, 2018). Chromatin accessibility landscapes can be used to detect a bias in differentiation trajectories. Greenleaf and colleagues

collected single cell chromatin accessibility profiles of defined human hematopoietic cell types. In line with recent mouse studies, this single cell analysis approach uncovered heterogeneity not only in the differentiated blood cell types but also in multipotent progenitors. The epigenetic state within progenitor cell populations correlated with patterns of hematopoietic differentiation that are suggestive of early lineage priming, a notion originally proposed by Tarik Moroy and colleagues (Zheng *et al.,* 2004) that challenges the current paradigm. In a more recent approach, researchers succeeded in the long-term *ex vivo* expansion of HSCs. Using clonally derived cultures, the authors uncovered considerable heterogeneity in the self-renewal capacity of HSCs (Wilkinson *et al.,* 2019). Furthermore, immunophenotypically specified cell types were found to 'flow' from one state to another, as an indication of plasticity (Buenrostro *et al.*, 2018). Taken together, these findings broaden the understanding of hematopoiesis, indicating that plasticity exist even around and within the HSC compartment.

CONCLUSION

From the examples given above, it is becoming clear that stem cell paradigms deduced from a single system should not simply be extrapolated to others. Tissues differ greatly in size and architecture. They exhibit widely divergent physiological functions, and are subject to diverse physical, chemical and biological insults. The HSC-paradigm has been applied as a stereotypic template to all adult stem cell biology, yet there is no good reason to assume that evolution has equipped each of our organs with the same 'one size fits all'- tissue maintenance strategy, rather than with 'smart', individualized stem cell strategies.

We propose to define adult tissue stem cell function operationally as the process by which a given tissue replaces its lost cells through cell division. If one looks at the above examples from this perspective, stem cell function may be embodied in HSC-like, "hard-wired", professional stem cells such as the HSC or the satellite cell of muscle. But stem cell function may also be executed in a 'diffuse' fashion by much larger populations of undifferentiated cells ('potential' stem cells), that neutrally compete for a limiting niche; good examples are the epidermis or the intestinal crypt. It may also be executed out by reserve stem cells: de-differentiated cells that are opportunistically formed from committed or even from fully differentiated cells upon tissue damage; the oval cell of the liver is a good example. And finally, tissues might avoid the involvement of undifferentiated stem cells altogether, by allowing differentiated cells to enter the cell cycle directly, such as illustrated by a regenerating liver upon partial resection. Soberingly, there may be no generalizable principles underlying the renewal mechanisms of the various individual tissues.

Phenotypic characteristics (such as the expression of known stem cell markers or the LRC state) can be taken as starting point for studies of an unexplored stem cell hierarchy. However, genetic tracing is rapidly becoming the method of choice for the essential, next step: establishing cellular lineage relationships. This can be done in mouse models by exploiting transgenic stem cell marker genes or using bar-coding strategies (Wu, Lee and Koo, 2019). By next-generation sequencing technologies, one can now also follow somatic mutation events within and between

human stem cell hierarchies (Lee-Six *et al.*, 2018). As an example of how the proliferative state of cells during tissue repair can be exploited to document tissue stem cell function, we have recently addressed the restorative response to myocardial infarction. A *Mki67*-RFP allele was used to identify all damage-induced proliferative cells; these were phenotyped by single cell mRNA sequencing. Simultaneously, a *Mki67*-IRES-CreERT2 allele was used to reveal all progeny of the proliferative cells by lineage tracing *in vivo*. This combined approach definitively identified the putative cKIT+ and Sca1+ cardiac stem cells as proliferative non-myocardial cells and established that the only reparative activity emanated from cardiac fibroblasts that generate a fibrotic scar. No significant *de novo* production of cardiomyocytes occurred from any of the cells that proliferated in response to the cardiac damage (Kretzschmar *et al.*, 2018). Thus, the technological breakthroughs of the past few years allow the study of tissue renewal and repair in unprecedented detail. Nature might still have some unanticipated stem cell surprises in store.

ACKNOWLEDGEMENT

DECLARATION OF INTEREST

H.C. is inventor on several patents related to organoid technology; his full disclosure is given at https://www.uu.nl/staff/JCClevers/.

REFERENCES

Barkauskas, C. E. *et al.* (2013) "Type 2 alveolar cells are stem cells in adult lung," *Journal of Clinical Investigation*, 123(7). doi: 10.1172/JCI68782.

Barker, N. *et al.* (2007) "Identification of stem cells in small intestine and colon by marker gene Lgr5," *Nature*, 449(7165). doi: 10.1038/nature06196.

Barrandon, Y. and Green, H. (1987) "Three clonal types of keratinocyte with different capacities for multiplication.," *Proceedings of the National Academy of Sciences of the United States of America*, 84(8).

Beltrami, A. P. *et al.* (2003) "Adult cardiac stem cells are multipotent and support myocardial regeneration.," *Cell*, 114(6).

van Berlo, J. H. *et al.* (2014) "c-kit+ cells minimally contribute cardiomyocytes to the heart," *Nature*, 509(7500). doi: 10.1038/nature13309.

van Berlo, J. H. *et al.* (2018) "van Berlo et al. reply," *Nature*, 555(7697). doi: 10.1038/nature25772.

Bickenbach, J. R. (1981) "Identification and behavior of label-retaining cells in oral mucosa and skin.," *Journal of dental research*, 60 Spec No C.

Bjerknes, M. and Cheng, H. (1999) "Clonal analysis of mouse intestinal epithelial progenitors.," *Gastroenterology*, 116(1).

Blokzijl, F. *et al.* (2016) "Tissue-specific mutation accumulation in human adult stem cells during life," *Nature*, 538(7624). doi: 10.1038/nature19768.

Brack, A. S. and Rando, T. A. (2012) "Tissue-Specific Stem Cells: Lessons from the Skeletal Muscle Satellite Cell," *Cell Stem Cell*, 10(5). doi: 10.1016/j.stem.2012.04.001.

Braun, K. M. *et al.* (2003) "Manipulation of stem cell proliferation and lineage commitment: visualisation of label-retaining cells in wholemounts of mouse epidermis," *Development*, 130(21). doi: 10.1242/dev.00703.

Braun, K. M. and Watt, F. M. (2004) "Epidermal Label-Retaining Cells: Background and Recent Applications," *Journal of Investigative Dermatology Symposium Proceedings*, 9(3). doi: 10.1111/j.1087-0024.2004.09313.x.

Breault, D. T. *et al.* (2008) "Generation of mTert-GFP mice as a model to identify and study tissue progenitor cells," *Proceedings of the National Academy of Sciences*, 105(30). doi: 10.1073/pnas.0804800105.

Buczacki, S. J. A. *et al.* (2013) "Intestinal label-retaining cells are secretory precursors expressing Lgr5," *Nature*, 495(7439). doi: 10.1038/nature11965.

Buenrostro, J. D. *et al.* (2018) "Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation," *Cell*, 173(6). doi: 10.1016/j.cell.2018.03.074.

Cai, C.-L. and Molkentin, J. D. (2017) "The Elusive Progenitor Cell in Cardiac Regeneration," *Circulation Research*, 120(2). doi: 10.1161/CIRCRESAHA.116.309710.

Cairns, J. (1975) "Mutation selection and the natural history of cancer," *Nature*. Nature Publishing Group, 255(5505). doi: 10.1038/255197a0.

Carrelha, J. *et al.* (2018) "Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells," *Nature*, 554(7690). doi: 10.1038/nature25455.

CHARGÉ, S. B. P. and RUDNICKI, M. A. (2004) "Cellular and Molecular Regulation of Muscle Regeneration," *Physiological Reviews*, 84(1). doi: 10.1152/physrev.00019.2003.

Cheng, H. and Leblond, C. P. (1974a) "Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine I. Columnar cell," *American Journal of Anatomy*, 141(4). doi: 10.1002/aja.1001410403.

Cheng, H. and Leblond, C. P. (1974b) "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(4). doi: 10.1002/aja.1001410407.

Cheshier, S. H. *et al.* (1999) "In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells.," *Proceedings of the National Academy of Sciences of the United States of America*, 96(6).

Chung, M.-I. *et al.* (2018) "Niche-mediated BMP/SMAD signaling regulates lung alveolar stem cell proliferation and differentiation," *Development*, 145(9). doi: 10.1242/dev.163014.

Clayton, E. *et al.* (2007) "A single type of progenitor cell maintains normal epidermis," *Nature*, 446(7132). doi: 10.1038/nature05574.

Comai, G. and Tajbakhsh, S. (2014) "Molecular and Cellular Regulation of Skeletal Myogenesis," in *Current topics in developmental biology*. doi: 10.1016/B978-0-12-405943-6.00001-4.

Cotsarelis, G., Sun, T. T. and Lavker, R. M. (1990) "Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis.," *Cell*, 61(7).

Crisan, M. and Dzierzak, E. (2016) "The many faces of hematopoietic stem cell heterogeneity," *Development*, 143(24). doi: 10.1242/dev.114231.

Decotto, E. and Spradling, A. C. (2005) "The Drosophila Ovarian and Testis Stem Cell Niches: Similar Somatic Stem Cells and Signals," *Developmental Cell*, 9(4). doi: 10.1016/j.devcel.2005.08.012.

Desai, T. J., Brownfield, D. G. and Krasnow, M. A. (2014) "Alveolar progenitor and stem cells in lung development, renewal and cancer," *Nature*, 507(7491). doi: 10.1038/nature12930.

Dollé, L. *et al.* (2015) "EpCAM and the biology of hepatic stem/progenitor cells," *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 308(4). doi: 10.1152/ajpgi.00069.2014.

Doupe, D. P. *et al.* (2012) "A Single Progenitor Population Switches Behavior to Maintain and Repair Esophageal Epithelium," *Science*, 337(6098). doi: 10.1126/science.1218835.

Ellison, G. M. *et al.* (2013) "Adult c-kitpos Cardiac Stem Cells Are Necessary and Sufficient for Functional Cardiac Regeneration and Repair," *Cell*, 154(4). doi: 10.1016/j.cell.2013.07.039.

van Es, J. H. *et al.* (2012) "Dll1+ secretory progenitor cells revert to stem cells upon crypt damage," *Nature Cell Biology*, 14(10). doi: 10.1038/ncb2581.

Escobar, M. *et al.* (2011) "Intestinal epithelial stem cells do not protect their genome by asymmetric chromosome segregation," *Nature Communications*. Nature Publishing Group, 2(1). doi: 10.1038/ncomms1260.

Evans, M. J. (2001) "The cultural mouse," *Nature Medicine*, 7(10). doi: 10.1038/nm1001-1081.

Evarts, R. P. *et al.* (1987) "A precursor-product relationship exists between oval cells and hepatocytes in rat liver.," *Carcinogenesis*, 8(11).

Finn, J. *et al.* (2019) "Dlk1-Mediated Temporal Regulation of Notch Signaling Is Required for Differentiation of Alveolar Type II to Type I Cells during Repair.," *Cell reports*. NIH Public Access, 26(11). doi: 10.1016/j.celrep.2019.02.046.

Font-Burgada, J. *et al.* (2015) "Hybrid Periportal Hepatocytes Regenerate the Injured Liver without Giving Rise to Cancer," *Cell*, 162(4). doi: 10.1016/j.cell.2015.07.026.

Foudi, A. *et al.* (2009) "Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells," *Nature Biotechnology*, 27(1). doi: 10.1038/nbt.1517.

Furuyama, K. *et al.* (2011) "Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine," *Nature Genetics*, 43(1). doi: 10.1038/ng.722.

Goldenring, J. R. and Mills, J. C. (2017) "Isthmus Time Is Here: Runx1 Identifies Mucosal Stem Cells in the Gastric Corpus," *Gastroenterology*, 152(1). doi: 10.1053/j.gastro.2016.11.028.

Grompe, M. (2014) "Liver Stem Cells, Where Art Thou?," *Cell Stem Cell*, 15(3). doi: 10.1016/j.stem.2014.08.004.

Hirsch, T. *et al.* (2017) "Regeneration of the entire human epidermis using transgenic stem cells," *Nature*. Nature Publishing Group, 551(7680). doi: 10.1038/nature24487.

Hu, H. *et al.* (2018). "Long-Term Expansion of Functional Mouse and Human Hepatocytes as 3D Organoids." *Cell*, 175(6), 1591-1606.e19.

Huch, M. *et al.* (2015). "Long-term culture of genome-stable bipotent stem cells from adult human liver." *Cell*, 160(1–2), 299–312.

Hsu, Y.-C., Pasolli, H. A. and Fuchs, E. (2011) "Dynamics between stem cells, niche, and progeny in the hair follicle.," *Cell*. NIH Public Access, 144(1). doi: 10.1016/j.cell.2010.11.049.

Jadhav, U. *et al.* (2017) "Dynamic Reorganization of Chromatin Accessibility Signatures during Dedifferentiation of Secretory Precursors into Lgr5+ Intestinal Stem Cells," *Cell Stem Cell*, 21(1). doi: 10.1016/j.stem.2017.05.001.

Jesty, S. A. *et al.* (2012) "c-kit+ precursors support postinfarction myogenesis in the neonatal, but not adult, heart," *Proceedings of the National Academy of Sciences*, 109(33). doi: 10.1073/pnas.1208114109.

Jones, P. H. and Watt, F. M. (1993) "Separation of human epidermal stem cells from transit amplifying cells on the basis of differences in integrin function and expression.," *Cell*, 73(4).

Klein, A. M. *et al.* (2010) "Mouse Germ Line Stem Cells Undergo Rapid and Stochastic Turnover," *Cell Stem Cell*, 7(2). doi: 10.1016/j.stem.2010.05.017.

Kretzschmar, K. *et al.* (2018) "Profiling proliferative cells and their progeny in damaged murine hearts," *Proceedings of the National Academy of Sciences*, 115(52). doi: 10.1073/pnas.1805829115.

Laurenti, E. and Göttgens, B. (2018) "From haematopoietic stem cells to complex differentiation landscapes," *Nature*. Nature Publishing Group, 553(7689). doi: 10.1038/nature25022.

Lee-Six, H. *et al.* (2018) "Population dynamics of normal human blood inferred from somatic mutations," *Nature*, 561(7724). doi: 10.1038/s41586-018-0497-0.

Leushacke, M. *et al.* (2013) "Lgr5+ Gastric Stem Cells Divide Symmetrically to Effect Epithelial Homeostasis in the Pylorus," *Cell Reports*, 5(2). doi: 10.1016/j.celrep.2013.09.025.

Leushacke, M. *et al.* (2017) "Lgr5-expressing chief cells drive epithelial regeneration and cancer in the oxyntic stomach," *Nature Cell Biology*, 19(7). doi: 10.1038/ncb3541.

Lin, S. *et al.* (2018) "Distributed hepatocytes expressing telomerase repopulate the liver in homeostasis and injury," *Nature*, 556(7700). doi: 10.1038/s41586-018-0004-7.

Lopez-Garcia, C. *et al.* (2010) "Intestinal Stem Cell Replacement Follows a Pattern of Neutral Drift," *Science*, 330(6005). doi: 10.1126/science.1196236.

MACKENZIE, J. C. (1969) "Ordered Structure of the Stratum Corneum of Mammalian Skin," *Nature*. Nature Publishing Group, 222(5196). doi: 10.1038/222881a0.

Michalopoulos, G. K. (2010) "Liver Regeneration after Partial Hepatectomy," *The American Journal of Pathology*, 176(1). doi: 10.2353/ajpath.2010.090675.

Miyajima, A., Tanaka, M. and Itoh, T. (2014) "Stem/Progenitor Cells in Liver Development, Homeostasis, Regeneration, and Reprogramming," *Cell Stem Cell*, 14(5). doi: 10.1016/j.stem.2014.04.010.

Morris, R. J. and Potten, C. S. (1994) "Slowly cycling (label-retaining) epidermal cells behave like clonogenic stem cells in vitro.," *Cell proliferation*, 27(5).

Morrison, S. J., Uchida, N. and Weissman, I. L. (1995) "The Biology of Hematopoietic Stem Cells," *Annual Review of Cell and Developmental Biology*, 11(1). doi: 10.1146/annurev.cb.11.110195.000343.

O'Connor, NicholasE. *et al.* (1981) "GRAFTING OF BURNS WITH CULTURED EPITHELIUM PREPARED FROM AUTOLOGOUS EPIDERMAL CELLS," *The Lancet*. Elsevier, 317(8211). doi: 10.1016/S0140-6736(81)90006-4.

Oransky, I. and Marcus, A. (2018) *Harvard and the Brigham call for more than 30 retractions of cardiac stem cell research*, Statnews.

Orkin, S. H. and Zon, L. I. (2008) "Hematopoiesis: An Evolving Paradigm for Stem Cell Biology," *Cell*, 132(4). doi: 10.1016/j.cell.2008.01.025.

Pardo-Saganta, A. *et al.* (2015) "Injury Induces Direct Lineage Segregation of Functionally Distinct Airway Basal Stem/Progenitor Cell Subpopulations," *Cell Stem Cell*, 16(2). doi: 10.1016/j.stem.2015.01.002.

Pfister, O. *et al.* (2008) "Role of the ATP-binding cassette transporter Abcg2 in the phenotype and function of cardiac side population cells.," *Circulation research*. NIH Public Access, 103(8). doi: 10.1161/CIRCRESAHA.108.174615.

Potten, C. S. (1974) "The epidermal proliferative unit: the possible role of the central basal cell.," *Cell and tissue kinetics*, 7(1).

Potten, C. S. *et al.* (1978) "The segregation of DNA in epithelial stem cells.," *Cell*, 15(3).

POTTEN, C. S. and MORRIS, R. J. (1988) "Epithelial stem cells in vivo," *Journal of Cell Science*. The Company of Biologists Ltd, 1988(Supplement 10). doi: 10.1242/jcs.1988.Supplement_10.4.

Potten, C. S., Owen, G. and Booth, D. (2002) "Intestinal stem cells protect their genome by selective segregation of template DNA strands.," *Journal of cell science*, 115(Pt 11).

Powell, A. E. *et al.* (2012) "The Pan-ErbB Negative Regulator Lrig1 Is an Intestinal Stem Cell Marker that Functions as a Tumor Suppressor," *Cell*, 149(1). doi: 10.1016/j.cell.2012.02.042.

1
2
3
4
5

Rama, P. *et al.* (2010) "Limbal Stem-Cell Therapy and Long-Term Corneal Regeneration," *New England Journal of Medicine*. Massachusetts Medical Society , 363(2). doi: 10.1056/NEJMoa0905955.

Raven, A. *et al.* (2017) "Cholangiocytes act as facultative liver stem cells during impaired hepatocyte regeneration," *Nature*, 547(7663). doi: 10.1038/nature23015.

Rawlins, E. L. *et al.* (2009) "The Role of Scgb1a1+ Clara Cells in the Long-Term Maintenance and Repair of Lung Airway, but Not Alveolar, Epithelium," *Cell Stem Cell*, 4(6). doi: 10.1016/j.stem.2009.04.002.

Rheinwald, J. G. and Green, H. (1975) "Formation of a keratinizing epithelium in culture by a cloned cell line derived from a teratoma.," *Cell*, 6(3).

Rodriguez-Fraticelli, A. E. *et al.* (2018) "Clonal analysis of lineage fate in native haematopoiesis," *Nature*, 553(7687). doi: 10.1038/nature25168.

Sacco, A. *et al.* (2008) "Self-renewal and expansion of single transplanted muscle stem cells," *Nature*, 456(7221). doi: 10.1038/nature07384.

Sangiorgi, E. and Capecchi, M. R. (2008) "Bmi1 is expressed in vivo in intestinal stem cells," *Nature Genetics*, 40(7). doi: 10.1038/ng.165.

Schaub, J. R. *et al.* (2014) "Evidence against a Stem Cell Origin of New Hepatocytes in a Common Mouse Model of Chronic Liver Injury," *Cell Reports*, 8(4). doi: 10.1016/j.celrep.2014.07.003.

Schepers, A. G. *et al.* (2011) "Lgr5 intestinal stem cells have high telomerase activity and randomly segregate their chromosomes," *The EMBO Journal*, 30(6). doi: 10.1038/emboj.2011.26.

Schofield, R. (1978) "The relationship between the spleen colony-forming cell and the haemopoietic stem cell.," *Blood cells*, 4(1–2).

Shimokawa, M. *et al.* (2017) "Visualization and targeting of LGR5+ human colon cancer stem cells," *Nature*. Nature Publishing Group, 545(7653). doi: 10.1038/nature22081.

Snippert, H. J. *et al.* (2010) "Intestinal Crypt Homeostasis Results from Neutral Competition between Symmetrically Dividing Lgr5 Stem Cells," *Cell*, 143(1). doi: 10.1016/j.cell.2010.09.016.

Stange, D. E. *et al.* (2013) "Differentiated Troy+ Chief Cells Act as Reserve Stem Cells to Generate All Lineages of the Stomach Epithelium," *Cell*, 155(2). doi: 10.1016/j.cell.2013.09.008.

Stanger, B. Z. (2015) "Cellular Homeostasis and Repair in the Mammalian Liver," *Annual Review of Physiology*, 77(1). doi: 10.1146/annurev-physiol-021113-170255.

Steinhauser, M. L. *et al.* (2012) "Multi-isotope imaging mass spectrometry quantifies stem cell division and metabolism," *Nature*, 481(7382). doi: 10.1038/nature10734.

Suda, T., Suda, J. and Ogawa, M. (1984) "Disparate differentiation in mouse hemopoietic colonies derived from paired progenitors.," *Proceedings of the National Academy of Sciences of the United States of America*, 81(8).

Takeda, N. *et al.* (2011) "Interconversion Between Intestinal Stem Cell Populations in Distinct Niches," *Science*, 334(6061). doi: 10.1126/science.1213214.

Tata, P. R. *et al.* (2013) "Dedifferentiation of committed epithelial cells into stem cells in vivo," *Nature*, 503(7475). doi: 10.1038/nature12777.

Tata, P. R. and Rajagopal, J. (2017) "Plasticity in the lung: making and breaking cell identity," *Development*, 144(5). doi: 10.1242/dev.143784.

Tetteh, P. W. *et al.* (2016) "Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters," *Cell Stem Cell*, 18(2). doi: 10.1016/j.stem.2016.01.001.

Tian, H. *et al.* (2011). "A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable". *Nature*, 478(7368), 255–259.

Uchida, S. *et al.* (2013) "Sca1-Derived Cells Are a Source of Myocardial Renewal in the Murine Adult Heart," *Stem Cell Reports*, 1(5). doi: 10.1016/j.stemcr.2013.09.004.

Vicinanza, C. *et al.* (2018) "Kitcre knock-in mice fail to fate-map cardiac stem cells," *Nature*, 555(7697). doi: 10.1038/nature25771.

Wang, B. *et al.* (2015) "Self-renewing diploid Axin2+ cells fuel homeostatic renewal of the liver," *Nature*, 524(7564). doi: 10.1038/nature14863.

White, R. B. *et al.* (2010) "Dynamics of muscle fibre growth during postnatal mouse development," *BMC Developmental Biology*. BioMed Central, 10(1). doi: 10.1186/1471-213X-10-21.

Wilkinson, A. C. *et al.* (2019). Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation. *Nature*.

Wilson, A. *et al.* (2008) "Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair," *Cell*, 135(6). doi: 10.1016/j.cell.2008.10.048.

Winton, D. J., Blount, M. A. and Ponder, B. A. J. (1988) "A clonal marker induced by mutation in mouse intestinal epithelium," *Nature*, 333(6172). doi: 10.1038/333463a0.

Wu, S.-H. S., Lee, J.-H. and Koo, B.-K. (2019) "Lineage Tracing: Computational Reconstruction Goes Beyond the Limit of Imaging.," *Molecules and cells*, 42(2). doi: 10.14348/molcells.2019.0006.

Yanger, K. *et al.* (2014) "Adult Hepatocytes Are Generated by Self-Duplication Rather than Stem Cell Differentiation," *Cell Stem Cell*, 15(3). doi: 10.1016/j.stem.2014.06.003.

Yellamilli, A. and van Berlo, J. H. (2016) "The Role of Cardiac Side Population Cells in Cardiac Regeneration," *Frontiers in Cell and Developmental Biology*, 4. doi: 10.3389/fcell.2016.00102.

Zeng, H. *et al.* (2004). "Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells," *The EMBO Journal*, 23(20), 4116–4125.

PROFILING PROLIFERATIVE CELLS AND THEIR PROGENY IN DAMAGED MURINE HEARTS

2

Kai Kretzschmar[a,b,1], Yorick Post[a,1], Marie Bannier-Hélaouët[a,b,c], Andrea Mattiotti[d], Jarno Drost[a,e], Onur Basak[a,f], Vivian S. W. Li[a,g], Maaike van den Born[a,b], Quinn D. Gunst[d], Danielle Versteeg[a], Lieneke Kooijman[a], Stefan van der Elst[a], Johan H. van Es[a,b], Eva van Rooij[a], Maurice J. B. van den Hoff[d], and Hans Clevers[a,b,e,2]

[a] Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences (KNAW) and University Medical Centre (UMC) Utrecht, 3584 CT Utrecht, The Netherlands
[b] Oncode Institute, Hubrecht Institute, Utrecht, 3584 CT Utrecht, The Netherlands
[c] École normale supérieure (ENS) de Lyon, 69007 Lyon, France
[d] Academic Medical Center, Department of Medical Physiology, 1105 AZ Amsterdam, The Netherlands
[e] Princess Máxima Center for Pediatric Oncology, 3584 CT Utrecht, The Netherlands
[f] UMC Utrecht, 3584 CG Utrecht, The Netherlands
[g] The Francis Crick Institute, London NW1 1AT, United Kingdom
[1] K.K. and Y.P. contributed equally to this work.
[2] To whom correspondence may be addressed. Email: h.clevers@hubrecht.eu

ABSTRACT

The significance of cardiac stem cell (CSC) populations for cardiac regeneration remains disputed. Here we apply the most direct definition of stem cell function (the ability to replace lost tissue through cell division) to interrogate the existence of CSCs. By single-cell messenger RNA sequencing and genetic lineage tracing using two Ki67 knock-in mouse models, we map all proliferating cells and their progeny in homoeostatic and regenerating murine hearts. Cycling cardiomyocytes were only robustly observed in the early postnatal growth phase, while cycling cells in homoeostatic and damaged adult myocardium represented various non-cardiomyocyte cell types. Proliferative post-damage fibroblasts expressing follistatin-like protein 1 (FSTL1) closely resemble neonatal cardiac fibroblasts and form the fibrotic scar. Genetic deletion of *Fstl1* in cardiac fibroblasts results in post-damage cardiac rupture. We find no evidence for the existence of a quiescent CSC population, for trans-differentiation of other cells types toward cardiomyocytes, or for proliferation of significant numbers of cardiomyocytes in response to cardiac injury.

Keywords

INTRODUCTION

Possible cell cycle re-entry of adult cardiomyocytes or derivation of new cardiomyocytes from putative resident progenitor cells upon myocardial infarction has been intensively investigated. Although with inconsistent results, these studies contributed to the general consensus that throughout adult mammalian life a small fraction of cardiomyocytes is replaced (Alkass et al. 2015; Soonpaa et al. 2015; Naqvi et al. 2014). Both the differentiation from a stem cell compartment and the division of pre-existing myocytes have been suggested as the source of new cardiomyocytes (Beltrami et al. 2003; Senyo et al. 2013). Nonetheless, the capacity of the adult mammalian heart to functionally regenerate upon injury remains controversial (van Berlo and Molkentin 2014; Uygur and Lee 2016; van Berlo et al. 2018; Vicinanza et al. 2018; Eschenhagen et al. 2017).

Studies on cardiac stem cells (CSCs) have relied heavily on specific stem cell markers that have been defined in unrelated stem cell systems such as c-KIT (Beltrami et al. 2003; Ellison et al. 2013), SCA-1 (Uchida et al. 2013) or ABCG2, a marker of so-called cardiac side population cells (Pfister et al. 2008; Yellamilli and van Berlo 2016). The validity of each of these markers has been disputed (van Berlo et al. 2014; Jesty et al. 2012; Cai and Molkentin 2017; van Berlo et al. 2018; Vicinanza et al. 2018). Also, anticipated changes in cellular metabolism or DNA label retention have been used to identify actively cycling cardiomyocytes, yet these methods do not allow visualisation of the cellular offspring (Kimura et al. 2015; Nakada et al. 2017). Furthermore, in these studies the proliferative capacity of resident non-cardiomyocyte cell lineages has been largely neglected. To address these issues, we sought to generate an unbiased map of proliferating cells and their progeny in neonatal, adult and post-damage murine hearts. Since the single defining characteristic of a stem cell is its ability to produce functional daughter cells by cell division (Clevers 2015), we posit that the most unbiased way of interrogating the involvement of stem cells in any biological growth or repair process is to genetically lineage trace all cells that proliferate during that pertinent biological process.

While absent from cells resting in the G_0 phase (Figure 1A), Ki67 is specifically expressed by all actively cycling cells in the cell-cycle stages of G_1, S, G_2 and M (Hutchins et al. 2010). It is therefore widely used in basic cell and developmental biology, in clinical oncology and in pathology as a general proliferation marker (Whitfield et al. 2006). We have generated *Mki67* knock-in mice (Basak et al. 2014; Basak et al. 2018) to perform genetic lineage tracing (Kretzschmar and Watt 2012) and comprehensively determine the offspring of any cell that becomes proliferative in neonatal, adult homoeostatic and damaged heart. With these genetic models, we set out to ask whether homoeostatic or damaged hearts harbour stem cells, under the premise that stem cells would have to enter the cell cycle to produce progeny replacing lost cells (Clevers 2015).

RESULTS

Cardiac cell proliferation is triggered in response to myocardial damage

We first aimed to determine the levels of cardiac cell proliferation in murine hearts at different postnatal time points and under different physiologic conditions. We therefore utilised a mouse model that expresses a red fluorescent protein (TagRFP) linked to the Ki67 protein (Basak et al. 2014) allowing for quantification and isolation of actively cycling live cells using flow cytometry (Figure 1B and *SI Appendix*, Figure S1). To analyse proliferating cardiac cells under homoeostatic conditions, we collected hearts of 1 and 8-week old *Mki67*[TagRFP] mice (**Figure 1C**). To evaluate cardiac cell proliferation upon injury, we subjected 8-week old *Mki67*[TagRFP] mice to a myocardial infarction (MI) by permanent ligation of the proximal left anterior descending (LAD) coronary artery and collected hearts 3, 7 and 14 days later (**Figure 1C**). The percentage of Ki67-RFP+ cells in the hearts of 1-week old neonatal mice was about 10%, while less than 0.05% of cells purified from the hearts of adult mice were positive for Ki67-RFP (**Figure 1D and E**). In mice subjected to MI surgery, the ratio of Ki67-RFP+ cells in the infarcted area was comparable to the corresponding region of hearts (apex) collected from sham-operated mice 3 days after the injury (**Figure 1D and E**). The percentage of Ki67-RFP+ cells doubled at 7 days post-MI and reached 1% enrichment at 14 days following the injury (**Figure 1D and E**). The percentage of Ki67-RFP+ cells in the remote tissue of both damaged and control hearts remained below 0.05% over the course of the experiment (**Figure 1E**).

We processed the purified Ki67-RFP+ cells for messenger RNA (mRNA) sequencing and analysed their gene-expression profiles using the DESeq2 algorithm (Love, Huber, and Anders 2014). Principal component analysis (PCA) showed very similar expression profiles of Ki67-RFP+ cells collected 3 and 7 days after MI (*SI Appendix*, **Figure S2**). Proliferative cells isolated from the hearts of neonatal or adult mice 14 days following injury, however, were distinctively different (*SI Appendix*, **Figure S2**). Neonatal Ki67-RFP+ cells were enriched for cardiomyocyte markers (*Myh6, Actc1, Tnnt2*), while Ki67-RFP+ cells from all other conditions showed enrichment for transcripts of non-cardiomyocyte markers (**Figure 1F**). Ki67-RFP+ cells collected within 1 week after injury were enriched for transcripts of haematopoietic genes such as *Ptprc* (encoding CD45), *Mrc1* and *Sirpa* suggesting the presence of a proliferative inflammatory infiltrate (**Figure 1F**). Genes specific for resident cell types such as endothelial cells (*Cdh5*, *Pecam1*, *Fabp4*) and cardiac fibroblasts (*Col1a1, Col1a2* and *Col6a1*) were more abundant in Ki67-RFP+ cells at 14 dpi (**Figure 1F**). Expression of the smooth muscle cell marker genes *Acta2*, *Actg2* and *Myh11* did not change upon damage in our dataset (**Figure 1F**).

Single-cell transcriptome analysis uncovers distinct proliferative populations within the murine heart

Fewer than 35% of cardiac ventricular cells are cardiomyocytes, the remainder belong to the endothelial, haematopoietic and fibroblast lineages (Pinto et al. 2016). A recent study has shown that single-cell transcriptome data can readily be generated from all cardiac cell lineages in the adult murine heart (Gladka et al. 2018). To generate a comprehensive transcriptome atlas
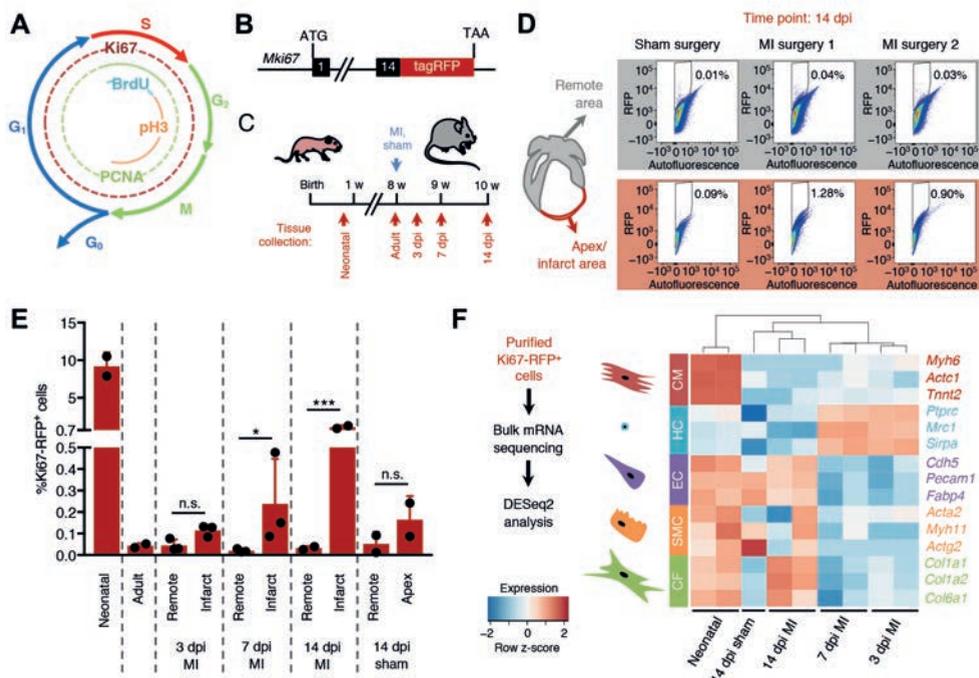
Figure 1. Quantification and characterisation of cardiac cell proliferation following injury. (A) Diagram showing the markers of the different phases of the cell cycle. (B) Schematic representation of the reporter mouse model expressing red fluorescent protein-tagged Ki67 (*Mki67*^TagRFP). (C) Experimental timeline of myocardial infarction (MI) surgery and tissue collection 3, 7 and 14 days post-MI injury (dpi). (D, E) Quantification of Ki67-RFP+ cells after MI or sham surgery in both remote (grey) and infarcted zone (red). The gating and sorting strategies are described in Figure S1. (D) Representative flow cytometry scatter plots 14 dpi. (E) Quantification of Ki67-RFP+ cells in neonatal, adult, infarct, remote and apex areas 3, 7 and 14 dpi (n = 2-3 mice per condition). All error bars represent ± SD. Asterisks indicate significance (Student's t test: n.s., not significant, p ³ 0.05; * $p < 0.05$; *** $p < 0.001$). (F) Bulk mRNA sequencing strategy and gene expression levels of different cardiac cell types identified in the purified Ki67-RFP+ populations. Cardiomyocytes were enriched in neonatal hearts. Haematopoietic cells were enriched 3 and 7 dpi, while resident lineages such as endothelial cells and fibroblasts were more abundant 14 dpi. CM, cardiomyocytes; HC, haematopoietic cells; EC, endothelial cells; SMC, smooth muscle cells; CF, cardiac fibroblasts.

of proliferative cells during postnatal growth, adult homoeostasis and cardiac regeneration, we therefore aimed to profile individual Ki67-RFP+ and Ki67-RFP- cardiac cells as well as unlabelled *Mki67*^wt/wt cardiac cells in a variety of settings: in 1-week old mice, when murine cardiomyocytes still have significant proliferative capacity (Alkass et al. 2015), in 8- to 10-week old mice when cardiomyocytes should be homoeostatic, and in adult mice after cardiac injury (**Figure 2A**). To collect cardiac cells from infarcted hearts, we chose the time point with the highest number of Ki67-RFP+ cells and at which most previous CSC studies have focused (Ellison et al. 2013; van Berlo et al. 2014; Wei et al. 2015): 14 days after MI surgery (**Figure 2A**). We performed single-cell mRNA-sequencing using the SORT-seq (sorting and robot-assisted transcriptome sequencing) method (Muraro et al. 2016) and analysed the sequencing data using the RaceID2 algorithm

A

*Mki67*tagRFP mouse or
CBA × C57BL/6 wildtype mouse

MI, I/R, sham

Birth | 1 w | 8 w | 9 w | 10 w

Tissue collection:

Neonatal | Adult | 14 dpi

B

(1) Heart dissection
(2) Cardiac cell isolation
(3) Single-cell sorting

Illumina P2
read 1
barcode
read 2

(4) Library preparation
(5) Sequencing
(6) Clustering

C Main cardiac cell subpopulations

t-SNE 2 / t-SNE 1

60 / 40 / 20 / 0 / −20 / −40 / −60

−60 −40 −20 0 20 40 60

Smooth muscle cells (SMC)

Endothelial cells (EC)

Cardiomyocytes (CM)

Haematopoietic cells (HC)

Cardiac fibroblasts (CF)

1,939 cells

D Number of cells of each cell subpopulation

SMC — 57 cells
HC — 458 cells
EC — 705 cells
CM — 437 cells
CF — 282 cells

0 200 400 600 800

E

Cardiac fibroblasts
Pdgfra, Col1a1, Tcf21

Cardiomyocytes
Myh6, Actc1, Tnnt2

Endothelial cells
Pecam1, Cdh5, Fabp4

Haematopoietic cells
Ptprc, Mrc1, Sirpa

Smooth muscle cells
Myh11, Acta2, Actg2

6 / 4 / 2 / 0 / −2

4 / 2 / 0 / −2

4 / 2 / 0 / −2

3 / 2 / 1 / 0 / −1 / −2 / −3

6 / 4 / 2 / 0 / −2

PDGFRα | MYH6 | PECAM1 (CD31) | CD45 | MYH11

F

S
G1
G2
M
G0

G

G0/G1

800 / 600 / 400 / 200 / 0

S

G2/M

Ki67-RFP+
Ki67-RFP−
unlabelled

***

40

(Grün et al. 2016) (**Figure 2B**). To exclude cells with low-quality RNA, we removed all cells expressing fewer than 1,000 transcripts. 1,939 cells passed this threshold, were clustered based on transcriptome similarities assessed by 1 – Pearson's correlation and grouped by *k*-medoids clustering (Grün et al. 2016). This *k*-medoids clustering compartmentalised the cells into 9 different major cell clusters, as visualised by *t*-distributed stochastic neighbour-embedding (*t*-SNE) (**Figure 2C** and *SI Appendix*, **Figure S3**). Based on differential gene expression of some well-characterised markers, we found that all major cardiac cell types were present in the single-cell dataset (**Figure 2C–E** and *SI Appendix*, **Figure S3**): clusters 8 and 9 comprised 437 cardiomyocytes enriched for transcripts of *Myh6*, *Actc1* and *Tnnt2,* clusters 6 and 7 contained a total of 282 cardiac fibroblasts expressing *Pdgfra*, *Col1a1* and *Tcf21*; 458 haematopoietic cells enriched for *Ptrpc*, *Mrc1* and *Sirpa* were assigned to clusters 1 and 4; 705 endothelial cells expressing *Pecam1, Cdh5* and *Fabp4* were present in clusters 2 and 3 and 57 cells in cluster 5 were enriched for smooth muscle genes *Acta2*, *Myh11* and *Actg2*. Of note, while *SIRPA* has been reported to be expressed by human embryonic stem cell-derived cardiac progenitors (Skelton et al. 2014), we only found significant *Sirpa* expression in haematopoietic cells in our dataset.

To validate that sequenced Ki67-RFP$^+$ cells were proliferative, we used the cyclone algorithm (Scialdone et al. 2015) to assign cell-cycle stages to each individual cell in our filtered dataset (**Figure 2F**). While present throughout the cell cycle (and absent from quiescent cells arrested in G_0) (Whitfield et al. 2006), Ki67 is most highly expressed in the G_2 stage (Bruno and Darzynkiewicz 1992; Bruno et al. 1991). Consistent with these previous studies, cells in S and G_2/M phase were enriched in the fraction of Ki67-RFP$^+$ cells (**Figure 2G**; Chi-squared test, $p < 0.001$), confirming that they were actively cycling.

Most Ki67-RFP$^+$ neonatal cells were identified as cardiomyocytes, while no Ki67-RFP$^+$ cardiomyocytes were found in the datasets compiled from either homoeostatic or injured adult hearts (*SI Appendix*, **Figure S4A**). In homoeostatic adult heart, Ki67-RFP$^+$ cells mapped to non-cardiomyocyte lineages, in particular to the haematopoietic and endothelial cell lineages (*SI*

◂ **Figure 2.** Single-cell transcriptome analysis uncovers distinct proliferative populations within the murine heart. (A) Experimental timeline for tissue collection of hearts from wildtype and *Mki67*^{RFP} mice, either neonatal or adults, 14 days after sham, ischaemia/reperfusion (I/R) or MI surgery (*n* = 2-4 mice per condition). (B) Schematic representation of SORT-seq workflow. Hearts were isolated (1) and digested into single-cell suspension (2), Ki67-RFP$^+$ and Ki67-RFP$^-$ cells were sorted into 384-well plates containing primers, dNTPs and spike-ins (3). Retrotranscription mix was distributed using Nanodrop II, material was pooled and amplified (4) before pair-end sequencing (5). Cells were clustered using RaceID2 (6). (C) Clustering of cardiac cells and cell-to-cell distances visualised by *t*-distributed stochastic neighbour-embedding (*t*-SNE) map, highlighting identified major cardiac cell-types. (D) Numbers of cells assigned to each cardiac cell lineage. (E) *t*-SNE map highlighting identified cell types based on previously described cellular markers (logarithmic scale of transcript expression). Markers expression is shown in lower panel by immunofluorescent staining. (F) *t*-SNE map displaying cell-cycle stage of each cell (S (red), G_2/M (green), G_0/G_1 (blue)) assigned by the cyclone algorithm. (G) *t*-SNE map showing the Ki67-RFP status from the flow cytometry data; Ki67-RFP$^+$ (red), Ki67-RFP$^-$ (black) or *Mki67*$^{wt/wt}$ cells without TagRFP construct (grey) and radar plot showing Ki67-RFP$^+$ cells enriched for the cycling G_2/M stage according to the cyclone algorithm. Asterisks indicate significance (Chi-square test: *** $p < 0.001$).

Appendix, **Figure S4A**). Ki67-RFP⁺ cells present in the infarcted region 14 dpi were mainly fibroblasts and haematopoietic cells (*SI Appendix*, **Figure S4A**), while Ki67-RFP⁺ cells in the remote region additionally represented mostly the endothelial and haematopoietic lineages (*SI Appendix*, **Figure S4A**). Collectively, these data show that there is a significant number of proliferative cardiomyocytes in neonatal hearts, but not in homoeostatic or injured adult hearts.

Since we did not find evidence for a distinct CSC population expressing proposed markers such as *Abcg2* (side population marker), *Kit* (encoding c-KIT) and *Ly6a* (encoding SCA-1), we investigated the presence of cells expressing at least one transcript of these markers within all cardiac cell types (**Figure 3A−C**). We identified 225 cells expressing *Abcg2*, which were scattered within all main cardiac cell types found in our dataset (**Figure 3A and B**) (Pfister et al. 2008; Yellamilli and van Berlo 2016). Most of the 145 *Kit*-expressing cells were endothelial cells (**Figure 3A and B**) (van Berlo et al. 2014), in contrast to earlier reports (Ellison et al. 2013; Beltrami et al. 2003). Some fibroblasts, which are usually not excluded when purifying putative c-KIT⁺ CSCs (Beltrami et al. 2003; Ellison et al. 2013; Vicinanza et al. 2017), also expressed *Kit* (**Figure 3A and B**). About 32% of all cells (613 cells of a total of 1,939 cells) in our dataset expressed *Ly6a* (**Figure 3A**). *Ly6a*-expressing cells were assigned to all cardiac cell types, while absent in neonatal cardiomyocytes (Uchida et al. 2013) (**Figure 3A and B**). Cells positive for a putative CSC marker (*i.e. Abcg2*, *Kit* or *Ly6a*) were identified in both Ki67-RFP⁺ and Ki67-RFP⁻ cardiac cell populations (**Figure 3C**). *Gata4*, a marker of regenerating cardiomyocytes in zebrafish (Kikuchi et al. 2010), was expressed in most cardiac lineages in a total of 214 cells (*SI Appendix*, **Figure S5**). The cardiomyocyte lineage marker *Nkx2-5* was only enriched in neonatal and adult cardiomyocytes (138 cells; *SI Appendix*, **Figure S5**), as shown before (Kasahara et al. 1998; Ranjbarvaziri et al. 2017). We further found expression of the foetal cardiac progenitor marker *Isl1* (Cai et al. 2003) in 87 cells belonging to the non-cardiomyocyte lineages (cardiac fibroblasts, endothelial cells and smooth muscle cells), but not in cardiomyocytes (*SI Appendix*, **Figure S5**). These data imply that cell culture, transplantation and lineage-tracing experiments relying on these putative CSC markers should be interpreted with great caution (Cai and Molkentin 2017).

To determine whether unique cell clusters appeared following MI in our single-cell mRNA-sequencing dataset, we quantified the number of cells from each experimental condition present within a given cluster (Dataset S1 and *SI Appendix*, **Figure S4B and C**). Cardiomyocytes clustered into two main clusters (clusters 8 and 9) irrespective of the cell-cycle stage (actively cycling vs. quiescent) (Dataset S1 and *SI Appendix*, **Figure S4B and C**). Endothelial cells clustered into clusters 2 and 3 irrespective of the experimental condition-of-origin (**Figure 2F**, Dataset S1 and *SI Appendix*, **Figure S4B and C**). The majority of haematopoietic cells in the dataset were purified from injured adult hearts (clusters 1 and 4), however, cluster 1 mainly contained haematopoietic cells present in homoeostatic (adult or neonatal) hearts (Dataset S1 and *SI Appendix*, **Figure S4B and C**). Fibroblast cluster 7 was predominantly composed of cells purified from homoeostatic adult hearts (65%). Surprisingly, while the majority of cells assigned to fibroblast cluster 6 were purified from injured adult heart (48%), 17% of the cells originated from neonatal hearts (*SI Appendix*, pie-chart inserts in **Figure S4C**). This suggests that in response

to ischaemic injury, some activated fibroblasts adopt a neonatal-like state. The fibroblasts in cluster 6 were predominantly Ki67-RFP[+] (**Figure 2F**, Dataset S1 and *SI Appendix*, **Figure S4B and C**) and had upregulated levels of genes involved in extracellular matrix remodelling such as *Cola1a2*, *Col3a1*, *Col5a1* and *Sparc* as well as genes encoding secreted proteins such as *Sfrp2* and *Dkk3* (Dataset S1).

To explore the cardiomyocyte fraction in more detail, we extracted all 437 cardiomyocytes out of the complete dataset and performed cell-type-specific subclustering. We re-ran the extracted cells through the RaceID2 algorithm with a higher cut-off of 2,000 transcripts per cell to allow for refined cell clustering. The 276 cardiomyocytes retained after quality filtering grouped into 6 *k*-medoids clusters (**Figure 3D**). Neonatal cardiomyocytes were enriched in clusters 1, 2 and 3, while adult injury-associated cardiomyocytes were mainly found in cluster 5 and adult homoeostatic cardiomyocytes in clusters 4 and 6 (**Figure 3D—F** and Dataset S2). We found robust expression of *Mki67* and other markers of proliferation (*Anln, Aurka, Birc5, Pcna* and *Plk1*) used by Soonpaa *et al.* (2015) in neonatal cardiomyocytes (**Figure 3G**), in agreement with our bulk mRNA-sequencing data (**Figure 1F**). However, expression of these proliferation markers was largely negligible in adult cardiomyocytes (**Figure 1F**). We performed differential gene-expression analysis by comparing neonatal and adult (homoeostatic and injury-associated) cardiomyocyte clusters to each other (**Figure 3H**). Among the 100 most upregulated genes in neonatal cardiomyocytes were genes related to the mitochondrial electron transport chain (*Cycs, Cox6a1*) and cardiac development (*Mef2a, Tnni1*) (**Figure 3I** and Dataset S3). Genes upregulated in adult homoeostatic cardiomyocytes included sarcomere assembly (*Tnni3*) and classical myosin marker (*Myh6*) (**Figure 3J** and Dataset S3). Adult injury-associated cardiomyocytes upregulated genes involved in oxidative damage response (*Gpx4*) and damage-induced cell cycle arrest (*Ccng1*) (**Figure 3K** and Dataset S3). Taken together, these data provide no evidence for the existence of a distinct CSC population in neonatal or adult (homoeostatic or post-damage) murine hearts and demonstrate that immature neonatal cardiomyocytes are readily proliferative. Yet, presence of proliferative cardiomyocytes in adult murine hearts remains low even after ischaemic injury, as indicated by the rare Ki67-RFP[+] cardiomyocyte found in our single-cell dataset collected from adult hearts (*SI Appendix*, **Figure S4A**).

Ki67 lineage tracing demonstrates *de novo* generation of cardiomyocytes in the neonatal murine heart

To identify the progeny of all proliferative cardiac cell types, we performed genetic lineage tracing of *Mki67*-expressing cells (Kretzschmar and Watt 2012). We used mice in which an internal ribosome entry site-Cre recombinase is fused to a tamoxifen-inducible version of the human oestrogen receptor (IRES-CreERT2) expression cassette, inserted downstream of the *Mki67* protein coding region (*Mki67*[IRES-CreERT2]) (Basak et al. 2018). We crossed these mice with Rosa26-CAG-loxP-stop-loxP-(LSL)-tdTomato reporter mice (hereafter LSL-tdTomato mice; **Figure 4A**) (Madisen et al. 2010). To assess initial labelling of proliferating *Mki67*-expressing cardiac cells, we first treated 1-week old (postnatal day 7, P7) *Mki67*[IRES-CreERT2] ×

A

All cells
1,939 cells

Abcg2-expressing cells
225 cells

Kit (c-KIT)-expressing cells
145 cells

Ly6a (SCA1)-expressing cells
613 cells

● Cardiac fibroblasts ● Cardiomyocytes ● Endothelial cells ● Haematopoietic cells ● Smooth muscle cells

B Proportion of cardiac cell types per marker

Ly6a
Kit
Abcg2
all

0   0.25   0.5   0.75   1

C Proportion of Ki67-RFP⁺ cells per marker

Ly6a
Kit
Abcg2
all

0   0.25   0.5   0.75   1

Ki67-RFP⁺
Ki67-RFP⁻
unlabelled

D CM subclustering
276 cells

E Transcriptome similarities

Cell cluster
1 – Pearson correlation
Main CM groups

F Main CM groups

● Neonatal CMs ● Adult injury-associated CMs ● Adult homeostatic CMs

G

Anln   *Aurka*
Birc5   *Mki67*
Pcna   *Plk1*

Transcript expression

Gene expression in single CMs

Neonatal
Adult injury-associated
Adult homeostatic

H TOP 100 upregulated genes in CM groups

Genes

Cells

Expression
−5 0 5
Row z-score

TOP 100 genes enriched in
■ Neonatal CMs
■ Adult injury-associated CMs
■ Adult homeostatic CMs

I

Cycs   *Cox6a1*
Mef2a   *Tnni1*

Transcript expression

J

Tnni3   *Myh6*

K

Gpx4   *Ccng1*

44

LSL-tdTomato mice with tamoxifen and collected tissue after 1 day (1 dpt) (**Figure 4B**). We found many tdTomato[+] cardiomyocytes 1 dpt (51.75 ± 0.77% of all tdTomato[+] cells, *i.e.* 568 of an estimated 4,000 cardiomyocytes analysed were tdTomato[+]; **Figure 4C and F**), confirming our single-cell transcriptome data. To trace the progeny of proliferating cells, we also collected hearts 1 week (7 dpt) and 2 months (2 mpt) after tamoxifen treatment at P7 (**Figure 4B**). At 7 dpt, 21.23 ± 2.32% of all tdTomato[+] cells were cardiomyocytes (**Figure 4D and F**), while tdTomato[+] cardiomyocytes represented 13.01 ± 0.87% of all reporter-labelled cells at 2 mpt (**Figure 4E and F**). This demonstrated that our approach also readily labels proliferative cardiomyocytes, but that the bulk of the cells produced from P7 proliferative (non-cardiomyocyte) cells long-term represented non-cardiomyocyte lineages.

Cardiac cells of non-cardiomyocyte lineages are continuously replenished during adult homoeostasis in murine hearts

To trace proliferating cardiac cells during adult homoeostasis, we next treated 8-week old *Mki67*[IRES-CreERT2] × LSL-tdTomato mice with tamoxifen and collected hearts 1 dpt and 1.5 year following tamoxifen exposure (1.5 ypt; **Figure 5A**). Reporter labelling in heart tissue was similar in mice injected 1 or 5 times with tamoxifen, with approximately 30 tdTomato[+] cells per 20× field (**Figure 5B and C**). In agreement with previous studies and a recent consensus statement describing modest cardiomyocyte turnover in adult mouse hearts (Alkass et al. 2015; Senyo et al. 2013; Eschenhagen et al. 2017), we found no tdTomato[+] cardiomyocytes 1 dpt and only very few tdTomato[+] cardiomyocytes 1.5 ypt (0.16% of all tdTomato[+] cells or 11 of an estimated 8 million cardiomyocytes analysed were tdTomato[+]; **Figure 5B and C**). The vast majority of proliferative cardiac cells labelled with tdTomato 1 day after tamoxifen injection were CD31[+] endothelial cells (**Figure 5D**), while the remaining tdTomato[+] cells were either PDGFRα[+] fibroblasts (**Figure 5E**) or CD45[+] haematopoietic cells (**Figure 5F**). We found a similar abundance of these cell types among all tdTomato[+] cells 1.5 ypt (**Figure 5D–F**), suggesting a continuous cellular turnover of these non-cardiomyocyte lineages in the adult homoeostatic murine heart.

◀ **Figure 3.** Expression patterns of putative cardiac stem cell markers and cardiomyocytes. (A) *t*-SNE map displaying stem cell marker (*Abcg2, Kit* and *Ly6a*) expressing cells in major cell type clusters; CM, cardiomyocytes; HC, haematopoietic cells; EC, endothelial cells; SMC, smooth muscle cells; CF, cardiac fibroblasts. (B) Quantification of marker-based cell fraction per cell type reveals most stem cell markers cover endothelial cells but can be found in all resident populations. (C) Quantification of marker-based cell fraction in sorted populations (unlabelled, Ki67-RFP[+] and Ki67-RFP[-]). (D) *t*-SNE map of cardiomyocyte subclusters (*n* = 6) identified using the RaceID2 algorithm. (E) Heatmap representation of the transcriptome similarities between cardiomyocytes from the 6 identified subclusters combined into 3 main groups of cardiomyocytes. (F) *t*-SNE map highlighting the cells assigned to each of the 3 main groups of cardiomyocyte subclusters. (G) Expression profile of representative proliferation markers in neonatal and adult (homoeostatic or injury-associated) cardiomyocytes. Gene expression is shown on the *y*-axis as transcript counts per cell on *x*-axis with the running mean in black. (H) Heatmap representation of top 100 upregulated genes in neonatal and adult homoeostatic or injury-associated) cardiomyocytes. (I, J, K) Violin plots showing the expression of representative genes upregulated in neonatal cardiomyocytes (I), adult homoeostatic cardiomyocytes (J), adult injury-associated cardiomyocytes (K).
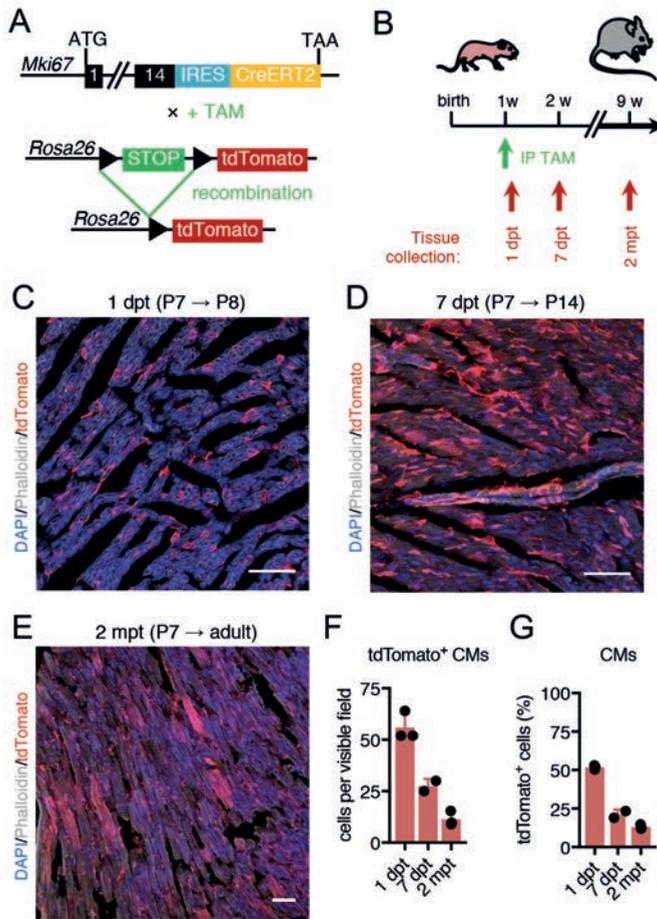
Figure 4. Ki67 lineage tracing demonstrates *de novo* generation of cardiomyocytes in the neonatal murine heart. (A) Genetic cross between *Mki67*[IRES-CreERT2] and LSL-tdTomato reporter mice to lineage trace *Mki67*-expressing cells. TAM, tamoxifen. (B) Neonatal hearts were collected 1 and 7 days as well as 8 weeks post tamoxifen (dpt) injection. Tamoxifen was injected at the age of 1 week. IP, intraperitoneal injection. (C, D, E) tdTomato-labelling 1 day (C), 7 days (D) and 2 months (E) after tamoxifen injection at postnatal day 7 (P7); months post tamoxifen (mpt). Nuclei were stained with DAPI (blue). Phalloidin (green or grey) was used to stain polymerised F-actin. Scale bars: 500 mm. (F) Percentage of tdTomato[+] cardiomyocytes of all tdTomato[+] cells at 1 dpt (P8; $n = 3$ mice), 7 dpt (P14; $n = 2$ mice) and 2 mpt ($n = 3$ mice). Asterisks indicate significance (Student's *t* test: * $p < 0.05$; *** $p < 0.001$).

Ki67 lineage tracing reveals proliferative response by non-cardiomyocyte cell lineages following ischaemic injury

It has been proposed that endogenous quiescent CSCs re-enter the cell cycle in response to cardiac injury and subsequently regenerate the myocardium (Ellison et al. 2013). To trace such proliferating cells following ischaemic injury, we performed MI surgery on
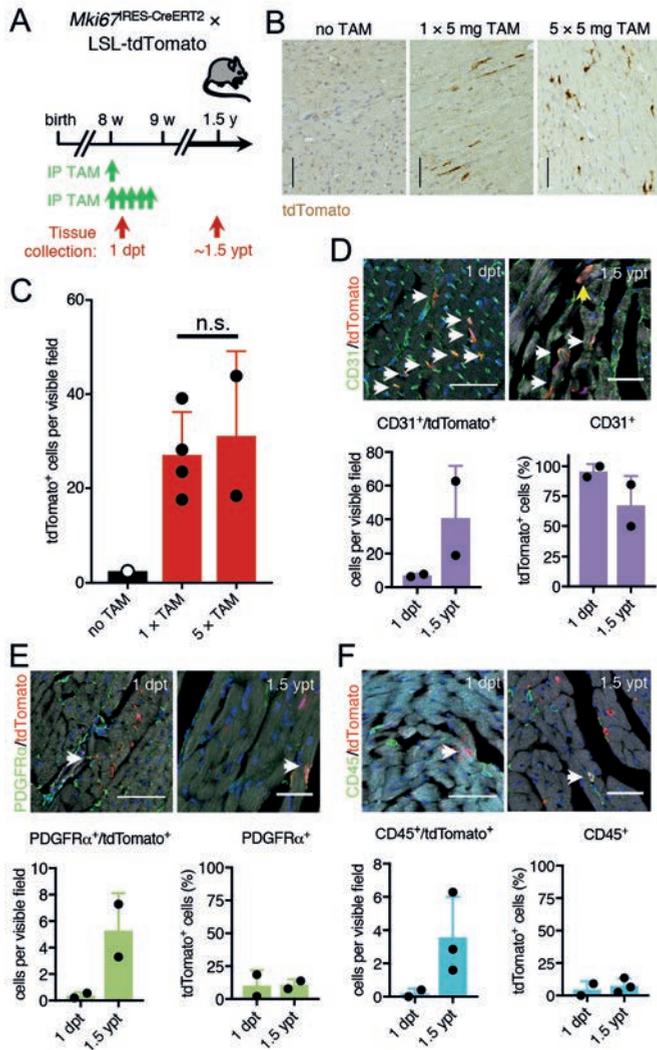
Figure 5. Continuous cellular turnover of non-cardiomyocyte lineages during adult homoeostasis of the murine heart. (A) Timeline of tamoxifen injection and tissue collection in adult *Mki67*[IRES-CreERT2] × LSL-tdTomato mice. (B, C) Representative images of stained paraffin sections (B) and corresponding quantification (C) of tdTomato labelling 1.5 year after tamoxifen exposure (1.5 ypt). Mice were either injected once with 5 mg tamoxifen (*n* = 4 mice) or 5 times with 5 mg tamoxifen (*n* = 2 mice). Non-injected mice were used as control (*n* = 1 mouse). Scale bars: 100 mm. (D-F) Representative images and quantification of co-stainings of tdTomato-traced cells (red) at 1 dpt and 1.5 ypt and CD31[+] (D), PDGFRα[+] (E) or CD45[+] cells (F) (green) (*n* = 2-3 mice), years post tamoxifen (ypt). White arrows point at double-positive cells. The yellow arrow shows one of the tdTomato-labelled cardiomyocytes we found across sections. Nuclei were counterstained with DAPI (blue). Phalloidin (grey) was used to stain polymerised F-actin. Scale bars: 50 mm (top panels) and 30 mm (bottom panels).

8-week old adult $Mki67^{\text{IRES-CreERT2}} \times$ LSL-tdTomato mice that received tamoxifen every other day for a total of 3 doses following the MI surgery (**Figure 6A**). We chose these injection time points to label proliferative cardiac cells in the acute phase of the injury response of the myocardium, when proliferation of CSCs and *de novo* generation of cardiomyocytes had been described in previous studies (Ellison et al. 2013; Smart et al. 2011). Sham control hearts displayed tdTomato-labelling similar to tracings found in homoeostatic adult hearts, while hearts collected 1 week following MI showed abundant tdTomato⁺ cells in the infarcted area (**Figure 6B and C**). One month following MI, scar tissue was almost entirely composed of tdTomato⁺ cells and robust tdTomato-labelling was present in the border zone (**Figure 6D and E**). No tdTomato⁺ cardiomyocytes were found in the infarct zone, and we found no significant difference in the percentage of tdTomato⁺ cardiomyocytes in the MI border zone compared to the remote area, or to the hearts of sham mice (**Figures 6F–I**; Student's *t* test, $p \geq 0.05$). Thus, ischaemic injury triggers a dramatic proliferative response in the infarcted myocardium including in the border zone, but does not stimulate significant generation of new cardiomyocytes.

FSTL1 expression is characteristic to a population of injury-activated cardiac fibroblasts

To more carefully document the role of cardiac fibroblasts in post-damage repair, we went back to our single-cell transcriptome dataset and performed refined cell-type-specific subclustering. We first extracted the cardiac fibroblast fraction from our initial analysis (257 cells) and re-ran these through the RaceID2 algorithm with a higher cut-off of 2,000 transcripts per cell to allow for refined cell clustering (as done for the cardiomyocyte subclustering). After filtering, the 243 remaining cardiac fibroblasts clustered into 11 subclusters (*SI Appendix*, **Figure S6A**). Cardiac fibroblasts purified from homoeostatic adult heart were mainly found in clusters 1, 3, 5 and 7, while neonatal fibroblasts and in the neonatal-like injury-associated fibroblasts were enriched in clusters 2, 4, 8–11 (*SI Appendix*, **Figure S6A and B**). We combined these clusters and performed differential gene-expression analysis (**Figure 7A**). Lineage markers *Pdgfra* and *Vim* were robustly expressed in all fibroblasts (**Figure 7B**) (Tallquist and Molkentin 2017; Pinto et al. 2016; Zeisberg and Kalluri 2010), while *Gsn*, a gene previously associated with cardiac remodelling after MI (Li et al. 2009), was upregulated in clusters enriched for adult homoeostatic fibroblasts (**Figure 7C** and Dataset S4). Fibroblasts also robustly expressed *Fstl1*, a marker that was recently described as a cardiomyogenic factor of epicardial origin expressed in the adult murine heart (Wei et al. 2015). Expression of *Fstl1* was significantly upregulated in the clusters containing neonatal fibroblasts and neonatal-like injury-associated fibroblasts, but *Fstl1* was also expressed by the fibroblasts in the healthy adult heart (**Figure 7C** and Dataset S4).

To further explore this observation, we utilised a mouse model carrying an *eGFP-IRES-CreERT2* expression cassette inserted into the *Fstl1* locus (Sylva et al. 2011) (**Figure 7D**). As suggested by our single-cell RNA-sequencing dataset, eGFP⁺ cells in the neonatal hearts of these mice were cardiac fibroblasts residing in the myocardium (*SI Appendix*, **Figure S6C–I**).

In contrast to previous reports (Wei et al. 2015), we did not find any eGFP expression in the epicardium (*SI Appendix*, **Figure S6H**), eGFP$^+$ cells in the adult heart were PDGFRα$^+$ cardiac fibroblasts (*SI Appendix*, **Figure S6J**).

Next, we asked whether *Fstl1*-eGFP$^+$ fibroblasts residing in the myocardium were a source for the fibroblast populations that appear in the infarcted area following cardiac ischaemic injury. We therefore generated *Fstl1*$^{eGFP-IRES-CreERT2}$ × LSL-tdTomato mice to enable genetic lineage tracing of FSTL1$^+$ cells (**Figure 7D**). To exclusively label cardiac fibroblasts residing in the homoeostatic myocardium and thereby avoid possible Cre-induced recombination in cells expressing *Fstl1* in response to ischaemic injury, we injected tamoxifen 1 week before induction of the MI and collected the tissue 2 weeks following surgery (**Figure 7E**). In mice undergoing sham surgery, no significant numbers of tdTomato$^+$ cells were found. We made similar observations in the remote myocardium of injured hearts (**Figure 7F and G**). However, robust numbers of tdTomato$^+$ fibroblasts were found in the infarcted region of injured hearts (**Figure 7F, H and I**) confirming that these cells are indeed derived from pre-existing cardiac fibroblasts infiltrating the infarct following cardiomyocyte apoptosis. Finally, we stained tissue sections from infarcted hearts with an antibody specific to FSTL1. As expected, confocal microscopy revealed that most FSTL1 immunoreactivity was present in the infarcted region of the myocardium and the border zone (*SI Appendix*, **Figure S6K–N**).

Collagens such as *Col1a1* and *Col3a1* were robustly expressed by cardiac fibroblasts (**Figure 7J**) and were strongly upregulated in response to cardiac injury (Dataset S4). Fibroblasts expressing these collagens made up a large fraction of the fibrotic scar (**Figure 7K**). To elucidate the functional role of FSTL1 in the formation of the fibrotic scar, we generated *Col1a1*-CreERT2 × *Fstl1*$^{flox/flox}$ mice to enable conditional knockout (cKO) of *Fstl1* in cardiac fibroblasts (**Figure 7L**) (Sylva et al. 2011). We injected cKO and control mice with tamoxifen 2 weeks prior to LAD occlusion to specifically remove *Fstl1* from the *Col1a1*-expressing fibroblast population. While loss of *Fstl1* did not seem to have an effect under homoeostatic conditions (**Figure 7M**), it resulted in a high mortality rate within the first 5 days following LAD occlusion with all cKO mice dying due to cardiac rupture, a rupture of the left ventricular-free wall (**Figure 7M**). Heterozygous cKO or wild-type control animals, with 1 or 2 remaining functional *Fstl1* alleles, had a significantly lower mortality rate (**Figure 7M**; log-rank test, $p = 0.0004$). Characterisation of the fibrotic scar in homozygous cKO mice demonstrated fewer Ki67$^+$ proliferating cells in both the border and the infarct zones compared to control (**Figure 7N**). Collectively, these data imply that FSTL1 is a marker of injury-activated fibroblasts that –in an autocrine fashion– signals to preventing cardiac rupture upon damage.

DISCUSSION

We have generated an exhaustive map of proliferative cardiac cells and their progeny during three critical conditions of postnatal life. Here we assessed cardiac cell proliferation in an unbiased fashion. Most studies on cardiac regeneration focus on cell cycle re-entry of cardiomyocytes or

A Mki67[IRES-CreERT2] × LSL-tdTomato

MI/IR/sham

Birth 8w 9w 10w 11w 12w

IP TAM

Tissue collection: 7 dpi   28 dpi

DAPI/autofluorescence (cardiomyocytes)/tdTomato

B **C**
sham control (28 dpi)

D **E**
MI surgery (7 dpi)

F **G**
MI surgery (28 dpi)

H

n.s.   n.s.

tdTomato+ CMs/20· image

1.0

0.5

0.0

infarct zone   border zone   remote/sham

tdTomato+ CMs (% of all)

0.15

0.10

0.05

0.00

infarct zone   border zone   remote/sham
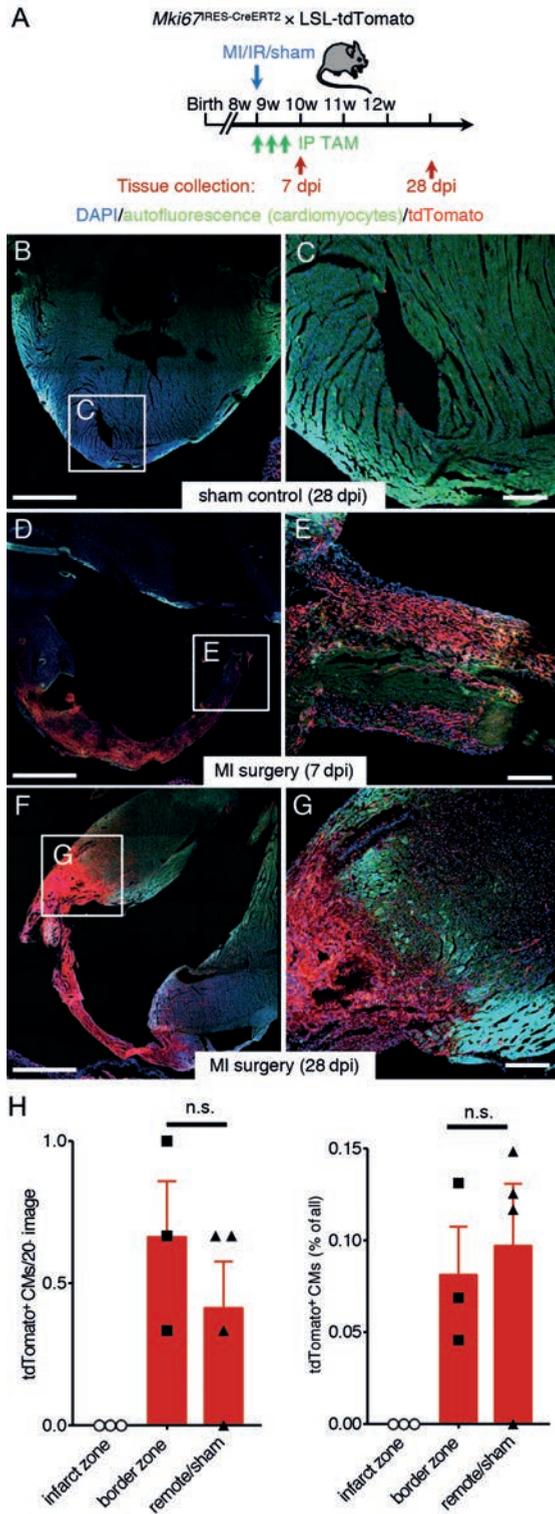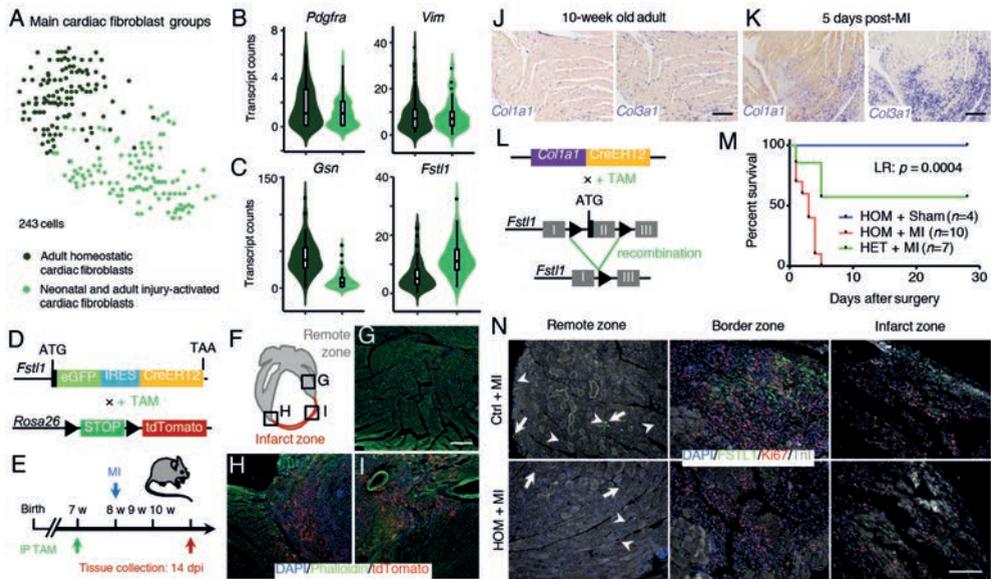
Figure 7. FSTL1 expression is characteristic to a population of injury-activated fibroblasts and is crucial in preventing cardiac rupture upon damage. (A) *t*-SNE map of cardiac fibroblast subclusters identified using the RaceID2 algorithm projecting the different experimental conditions. (B, C) Violin plots showing the expression of *Pdgfra*, *Vim* (B), *Gsn* and *Fstl1* (C) in neonatal and injury-activated fibroblasts (light green) and homoeostatic adult fibroblasts (dark green). (D) Schematic representation of the generation of mice expressing eGFP and CreERT2 by cassette insertion within the *Fstl1* protein coding region, enabling lineage tracing by crossing it with LSL-tdTomato mice. (E) Experimental timeline. (F-I) Progeny (red) of *Fstl1*-expressing fibroblasts is enriched in the scar tissue 14 dpi (H-I), while tdTomato-labelling in the remote area remains low (G). Nuclei were counterstained with DAPI (blue) and F-actin visualised by Phalloidin (green). (J, K) *In situ* hybridisation of *Col1a1* and *Col3a1* in adult homoeostatic heart (J) and 5 days after myocardial infarction (K). (L) Schematic representation of the generation of conditional knockout (cKO) mice by crossing *Col1a1*-CreERT2 mice with *Fstl1*flox/flox mice. Resulting in the depletion of *Fstl1* exon2 upon tamoxifen induction. (M) Kaplan-Meier survival curve of homozygous cKO mice after MI (red), homozygous cKO mice after sham surgery (blue) and heterozygous cKO mice after MI (green) over the course of 30 days. (N) Immunofluorescence stainings of FSTL1 (green), Ki67 (red) and TnI (white) in damaged control and homozygous cKO hearts in remote, border and infarct zone. Nuclei were counterstained with DAPI (blue). Scale bars: 100 mm.

◄ **Figure 6.** Ki67 lineage tracing reveals proliferative response following ischaemic injury. (A) Experimental timeline. (B, C) Visualisation of rare tdTomato+ cells in cryosections of sham control mice 28 dpi (*n* = 1 mouse). (D-G) Large accumulation of tdTomato+ cells in the infarct zone 7 dpi (D-E, *n* = 2 mice) and 28 dpi (F-G, *n* = 3 mice). The myocardium was visualised using cardiomyocyte autofluorescence; CM (green), nuclei were counterstained with DAPI (blue). (H) Quantification of tdTomato+ cardiomyocytes (CMs) per section scored (*n* = 3-4 mice). Student's *t* test: n.s., not significant. (I) Percentage of tdTomato+ cardiomyocytes (*n* = 3-4 mice). Student's *t* test: n.s. Scale bars: 1 mm (B, D, F) and 200 mm (C, E, G).

the existence of resident stem cells, thereby restricting the analysis by the use of generic stem cell markers or of DNA label retention as a surrogate stem cell marker (Beltrami et al. 2003; Ellison et al. 2013; Uchida et al. 2013; Senyo et al. 2013; Kimura et al. 2015). Hence, we think that our dataset allows resolution of current controversies on cardiac regeneration (Cai and Molkentin 2017; Uygur and Lee 2016; van Berlo and Molkentin 2014; van Berlo et al. 2018; Vicinanza et al. 2018).

While our data demonstrate that neonatal murine cardiomyocytes are actively cycling (**Figure 1 and 2**), we did not find evidence for significant *de novo* generation of cardiomyocytes in the homoeostatic or regenerating adult murine heart through cell division (**Figure 4 and 5**). In accordance with this, recent studies by the laboratories of Jonas Frisén and Olaf Bergmann used radiocarbon dating to demonstrate that the number of cardiomyocytes in human and mice is mainly established in the perinatal period and turnover rates decrease throughout life (Alkass et al. 2015; Bergmann et al. 2015). These data were confirmed by other means (Soonpaa et al. 2015; Kimura et al. 2015). Of note, significant induction of proliferation of adult post-mitotic cardiomyocytes has only been achieved *in vivo* by overexpressing cell-cycle regulators that are usually solely expressed by foetal cardiomyocytes (Mohamed et al. 2018).

It has been proposed that endogenous quiescent CSCs re-enter the cell cycle in response to cardiac injury (Ellison et al. 2013; Vicinanza et al. 2017; Vicinanza et al. 2018). We find no evidence for such CSCs that would contribute to the myocardial lineage upon damage. Further, we find that proposed CSC markers are widely expressed by non-cardiomyocyte cardiac lineages (**Figure 3**). Putative c-KIT+ CSCs, for example, are often purified by removing CD45+ haematopoietic cells and CD31+ endothelial cells (Beltrami et al. 2003; Ellison et al. 2013; Vicinanza et al. 2017). However, cardiac fibroblasts in our single-cell transcriptome dataset also express *Kit* transcripts. This demonstrates that *Kit* is expressed by a wide range of cardiac cell populations, which may explain some of the opposing results published on the role c-KIT+ cells for cardiac regeneration (Beltrami et al. 2003; Ellison et al. 2013; Vicinanza et al. 2017; van Berlo et al. 2014; van Berlo et al. 2018; Vicinanza et al. 2018).

In contrast to previous studies utilising label incorporation (*i.e.* BrdU) into the DNA of S-phase cells to assess proliferation of putative CSCs or cardiomyocytes (Ellison et al. 2013; Smart et al. 2011; Beltrami et al. 2003), we utilise genetic lineage tracing of Ki67+ cardiac cells to investigate evidence for *de novo* generation of cardiomyocytes in response to myocardial damage. Our data does not reveal significant numbers of newly formed cardiomyocytes (**Figure 5**), in line with a recent study assessing cardiomyocyte expansion during murine heart injury (Sereti et al. 2018). While we did not find evidence for cardiac regeneration through actively cycling cardiomyocytes or CSCs, we charted the robust injury response in the adult mammalian heart that essentially consists of proliferation of non-cardiomyocyte cardiac cell lineages (**Figure 1, 2 and 5**). The majority of proliferative cells in the adult homoeostatic heart are CD31+ cells of the vasculature (**Figure 5**). However, haematopoietic cells as well as cardiac fibroblasts robustly enter the cell cycle in the infarcted areas of the myocardium of injured adult mice (**Figure 6 and 7**). Cardiac fibroblasts proliferated in response to ischaemic injury and adopt a gene-expression profile similar to that of neonatal cardiac fibroblasts. This phenomenon

has previously been described for adult murine skin epidermis, where dermal fibroblasts are reprogrammed to a neonatal-like state in response to sustained activation of epidermal Wnt/β-catenin signalling (Collins, Kretzschmar, and Watt 2011). A subsequent study showed that adult fibroblast remodelling significantly improved epidermal regeneration upon wounding of adult murine skin (Driskell et al. 2013).

Our single-cell transcriptome dataset further showed that the proposed epicardial-specific cardiomyogenic factor FSTL1 (Wei et al. 2015) is prominently expressed by the cardiac fibroblast lineage. We do not find evidence for significant epicardial expression of FSTL1 under any of the conditions analysed. We confirmed our single-cell data using a *Fstl1*-eGFP reporter mouse as well as by histological staining (**Figure 7** and *SI Appendix*, **Figure S6**) (Sylva et al. 2011). Using genetic lineage tracing, we subsequently demonstrate that *Fstl1*-expressing fibroblasts contribute principally to the myocardial remodelling occurring in response to cardiac injury. Finally, we show that the genetic deletion of *Fstl1* in myocardial fibroblasts results in a severe phenotype with high mortality due to cardiac rupture upon injury (**Figure 7M**), a phenomenon similar to that described in another study with a less pronounced phenotype (Maruyama et al. 2016) . This demonstrates the importance of autocrine FSTL1 signalling in scar formation.

Overall, our results imply that all markers used in previous studies to identify CSCs lack appropriate specificity in the heart. In addition, we find that post-infarct generation of new cardiomyocytes through proliferation of any type of cardiac cells is negligible, when compared to the proliferative response of non-cardiomyocytes. We find proliferating cells within all other cardiac-cell lineages, suggesting that these lineages are not maintained by distinct stem cells, but rather through cell division within the cell lineage. The proliferative response of cardiac fibroblasts in particular is important for infarct healing, yet it does not generate new cardiomyocytes, but rather a post-infarction fibrotic scar.

METHODS
Mice

Generation and genotyping of *Mki67*[IRES-CreERT2] (Basak et al. 2018), *Mki67*[TagRFP] (Basak et al. 2014), *Fstl1*[eGFP-IRES-CreERT2] and *Fstl1*[flox/flox] (Sylva et al. 2011). LSL-tdTomato (B6.Cg-Gt(ROSA)26Sor[tm9(CAG-tdTomato)Hze]/J) (Madisen et al. 2010) and Col1a1-CreERT2 (B6.Cg-Tg(Col1a1[cre/ERT2])1Crm/J) (Kim, Nakashima, and de Crombrugghe 2004) were described elsewhere. All alleles were maintained on a mixed C57BL/6 background. All mouse experiments were conducted under a project license granted by the Central Committee Animal Experimentation (CCD) of the Dutch government and approved by the Hubrecht Institute Animal Welfare Body (IvD) and by the Animal Experimental Committee of the Academic Medical Center, Amsterdam. Littermates were used as controls, as indicated for each experiment. Both male and female mice were used in all experiments, except for mRNA-sequencing experiments, which were preferentially performed with male mice.

Surgical Procedures

Myocardial infarction (MI) was induced by permanent occlusion of the left anterior descending coronary artery (LAD). Ischaemia/reperfusion (I/R) was performed by LAD occlusion for 1 hour (ischaemic period), followed by LAD reperfusion. Detailed surgical procedures are described in the *SI Appendix, Materials and Methods*.

Genetic Lineage-Tracing Experiments

Genetics lineage tracing of *Mki67*-expressing cells or *Fstl1*-expressing cells was induced by treating, respectively, *Mki67*[IRES-CreERT2] × LSL-tdTomato or *Fstl1*[eGFP-IRES-CreERT2] × LSL-tdTomato with tamoxifen (Sigma-Aldrich) dissolved in sunflower oil via the IP route. Details on the experimental procedures, treatment regime and tamoxifen dosage injected are provided in the *SI Appendix, Materials and Methods*.

Conditional Knockout Experiment

For conditional knockout of *Fstl1* in cardiac fibroblasts 8-week-old Col1a1[CreERT2] × *Fstl1*[flox/flox] mice were injected with 10 mg/kg tamoxifen for 4 consecutive days. Two weeks after the start of tamoxifen injection, mice underwent permanent ligation of the left anterior descending (LAD) artery. Survival rate was monitored or mice were sedated with CO_2 and sacrificed up to 30 days after surgery.

Cardiac Cell Isolation

Mouse hearts were dissected, the atria were removed and, when suitable, hearts were divided up into remote and apex (*i.e.* for MI or I/R surgery, borderzone and infarcted zone). Whole ventricles were prepped from dissected hearts of neonatal mice. Tissue dissection was either performed as recently described (Gladka et al. 2018) or as detailed in the *SI Appendix, Materials and Methods*.

Flow Cytometric Purification

4',6-diamidino-2-phenylindole (DAPI) was added immediately prior to flow sorting. DAPI-negative and MitoTracker-positive living cells were either sorted into TRIzol reagent (Thermo Scientific) for bulk mRNA-sequencing or into 384-well plates containing 96 or 384 unique molecular identifier (UMI) barcode primer-sets, ERCC92 spike-ins (Agilent) and dNTPs (Promega) for single-cell mRNA-sequencing (SORT-seq) (Muraro et al. 2016) using a flow sorter (FACSAriaII, FACSFusion or FACSJazz, all BD). Sorted samples were stored at −80 °C until further processing.

Messenger RNA Sequencing

Samples were lysed and RNA from each bulk sort sample or single cell was barcoded and processed using the CEL-Seq2 technique (Hashimshony et al. 2016; Grün et al. 2016; Muraro

et al. 2016). Details on the sequencings performed are described in the *SI Appendix, Materials and Methods*. Bioinformatics analysis of sequenced libraries was performed as described below.

Histology

Histology on cryosections and Paraffin sections was performed using standard methods as detailed in the in the *SI Appendix, Materials and Methods*.

The following antibodies or labelling agents were used for immunostainings: rat anti-CD45 (30F11, eBiosciences), goat anti-mFSTL1 (R&D Systems, AF1738), goat anti-mouse PDGFRα (R&D Systems, AF1062), Phalloidin Alexa Fluor 488 (Life Technologies, A12379), Phalloidin Atto 647N (Sigma), rabbit anti-RFP/tdTomato (Rockland, 600-401-379), mouse anti-Ki67 (BD Bioscience, 556003), Alexa Fluor 568 donkey anti-rabbit IgG (Life Technologies, A10042) and Alexa Fluor 568 donkey anti-goat IgG (Life Technologies, A11057).

In Situ Hybridisation

In situ hybridisation (ISH) was performed on 8 μm paraffin sections of mouse hearts, as described (Sylva et al. 2011). Details ISH procedures are described in the *SI Appendix, Materials and Methods*.

Imaging

Images of cryosections were acquired on an inverted confocal laser scanning microscope (Leica SP8X and SP8). Paraffin sections were imaged using a Leica DM4000 or DM6000 optical microscope. ISH sections were imaged using a Leica DM5000 microscope.

Bioinformatics Analysis

DNA library sequencing, mapping to the mouse reference genome and quantification of transcript abundance were performed as described elsewhere (Muraro et al. 2016). Bulk sequencing libraries were analysed using the DESeq2 package (Love, Huber, and Anders 2014). Cell clusters and gene expression levels across clusters were generated using the RaceID2 algorithm (Grün et al. 2016) from single-cell libraries. All bioinformatics analysis was performed using R version 3.4.0 (R Foundation, https://www.r-project.org) and RStudio version 1.0.143 (https://www.rstudio.com). Detailed bioinformatics analysis is described in the *SI Appendix, Materials and Methods*.

Statistics

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to the sample allocation during experiments and outcome assessment. No animals were excluded from analysis. All data are presented as mean ± standard deviation (SD), unless stated otherwise. Statistically significant differences between the mean of two conditions or experimental groups were determined using an unpaired two-tailed Student's *t*-test. A Chi-square test (χ^2) was performed to determine statistical significance

for the data presented in Figure 2F. A log-rank (Mantel-Cox) test was performed to determine statistical significance for the data presented in Figure 7M. Data was obtained from at least two biological replicates per condition and from at least two independent experiments to ensure reproducibility. Kaplan-Meier survival curves were generated and all statistical analyses were performed using GraphPad Prism version 7.0 software.

Data Resources

Differentially expressed genes of the initial cardiac-cell clustering, of the cardiomyocyte subclustering, of the main cardiomyocyte clusters and of the cardiac fibroblast subclustering are provided in Datasets S1–S4, respectively. RNA-sequencing data are deposited in the Gene Expression Omnibus (GEO), accession number GSE102048, except for some single-cell transcriptome data of cardiac cells isolated from wildtype ($Mki67^{wt/wt}$) mice, which were published elsewhere (Gladka et al. 2018). All source data supporting the findings of this study are available from the corresponding author on request.

AUTHOR CONTRIBUTIONS

K.K., Y.P., J.D., E.v.R. and H.C. conceived the project. K.K., Y.P., J.D., A.M., M.B.v.d.H. and H.C. designed experiments and interpreted results. K.K., M.v.d.B., D.V., A.M., and L.K. performed animal experiments. J.H.v.E. and E.v.R. supervised animal experiments. K.K., Y.P., M.B.-H. and J.D. performed all histology and imaging experiments. S.v.d.E. performed flow cytometry. K.K., Y.P. and M.B.-H. performed sequencing experiments with initial assistance of O.B. K.K. performed bioinformatics analysis. O.B. generated the $Mki67^{TagRFP}$ and $Mki67^{IRES-CreERT2}$ mouse models. V.S.L. generated the $Fstl1^{eGFP-IRES-CreERT2}$ mouse model. A.M., Q.D.G. and M.B.v.d.H. performed experiments with the $Col1a1$-CreERT2 x $Fstl1^{flox/flox}$ mouse model. K.K., E.v.R. and H.C. acquired funding. K.K., Y.P., M.B.-H. and H.C. wrote the manuscript with input from all other authors.

ACKNOWLEDGEMENTS

REFERENCES

Alkass, K., J. Panula, M. Westman, T. D. Wu, J. L. Guerquin-Kern, and O. Bergmann. 2015. 'No Evidence for Cardiomyocyte Number Expansion in Preadolescent Mice', *Cell*, 163: 1026-36.

Basak, O., T. G. Krieger, M. J. Muraro, K. Wiebrands, D. E. Stange, J. Frias-Aldeguer, N. C. Rivron, M. van de Wetering, J. H. van Es, A. van Oudenaarden, B. D. Simons, and H. Clevers. 2018. 'Troy+ brain stem cells cycle through quiescence and regulate their number by sensing niche occupancy', *Proc Natl Acad Sci U S A*, 115: E610-E19.

Basak, O., M. van de Born, J. Korving, J. Beumer, S. van der Elst, J. H. van Es, and H. Clevers. 2014. 'Mapping early fate determination in Lgr5+ crypt stem cells using a novel Ki67-RFP allele', *EMBO J*, 33: 2057-68.

Beltrami, A. P., L. Barlucchi, D. Torella, M. Baker, F. Limana, S. Chimenti, H. Kasahara, M. Rota, E. Musso, K. Urbanek, A. Leri, J. Kajstura, B. Nadal-Ginard, and P. Anversa. 2003. 'Adult cardiac stem cells are multipotent and support myocardial regeneration', *Cell*, 114: 763-76.

Bergmann, O., S. Zdunek, A. Felker, M. Salehpour, K. Alkass, S. Bernard, S. L. Sjostrom, M. Szewczykowska, T. Jackowska, C. Dos Remedios, T. Malm, M. Andra, R. Jashari, J. R. Nyengaard, G. Possnert, S. Jovinge, H. Druid, and J. Frisen. 2015. 'Dynamics of Cell Generation and Turnover in the Human Heart', *Cell*, 161: 1566-75.

Bruno, S., H. A. Crissman, K. D. Bauer, and Z. Darzynkiewicz. 1991. 'Changes in cell nuclei during S phase: progressive chromatin condensation and altered expression of the proliferation-associated nuclear proteins Ki-67, cyclin (PCNA), p105, and p34', *Exp Cell Res*, 196: 99-106.

Bruno, S., and Z. Darzynkiewicz. 1992. 'Cell cycle dependent expression and stability of the nuclear protein detected by Ki-67 antibody in HL-60 cells', *Cell Prolif*, 25: 31-40.

Cai, C. L., X. Liang, Y. Shi, P. H. Chu, S. L. Pfaff, J. Chen, and S. Evans. 2003. 'Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart', *Dev Cell*, 5: 877-89.

Cai, C. L., and J. D. Molkentin. 2017. 'The Elusive Progenitor Cell in Cardiac Regeneration: Slip Slidin' Away', *Circ Res*, 120: 400-06.

Clevers, H. 2015. 'STEM CELLS. What is an adult stem cell?', *Science*, 350: 1319-20.

Collins, C. A., K. Kretzschmar, and F. M. Watt. 2011. 'Reprogramming adult dermis to a neonatal state through epidermal activation of beta-catenin', *Development*, 138: 5189-99.

Driskell, R. R., B. M. Lichtenberger, E. Hoste, K. Kretzschmar, B. D. Simons, M. Charalambous, S. R. Ferron, Y. Herault, G. Pavlovic, A. C. Ferguson-Smith, and F. M. Watt. 2013. 'Distinct fibroblast lineages determine dermal architecture in skin development and repair', *Nature*, 504: 277-81.

Ellison, G. M., C. Vicinanza, A. J. Smith, I. Aquila, A. Leone, C. D. Waring, B. J. Henning, G. G. Stirparo, R. Papait, M. Scarfo, V. Agosti, G. Viglietto, G. Condorelli, C. Indolfi, S. Ottolenghi, D. Torella, and B. Nadal-Ginard. 2013. 'Adult c-kit(pos) cardiac stem cells are necessary and sufficient for functional cardiac regeneration and repair', *Cell*, 154: 827-42.

Eschenhagen, T., R. Bolli, T. Braun, L. J. Field, B. K. Fleischmann, J. Frisen, M. Giacca, J. M. Hare, S. Houser, R. T. Lee, E. Marban, J. F. Martin, J. D. Molkentin, C. E. Murry, P. R. Riley, P. Ruiz-Lozano, H. A. Sadek, M. A. Sussman, and J. A. Hill. 2017. 'Cardiomyocyte Regeneration: A Consensus Statement', *Circulation*, 136: 680-86.

Gladka, M. M., B. Molenaar, H. de Ruiter, S. van der Elst, H. Tsui, D. Versteeg, G. P. A. Lacraz, M. M. H. Huibers, A. van Oudenaarden, and E. van Rooij. 2018. 'Single-Cell Sequencing of the Healthy and Diseased Heart Reveals Ckap4 as a New Modulator of Fibroblasts Activation', *Circulation*.

Grün, D., M. J. Muraro, J. C. Boisset, K. Wiebrands, A. Lyubimova, G. Dharmadhikari, M. van den Born, J. van Es, E. Jansen, H. Clevers, E. J. de Koning, and A. van Oudenaarden. 2016. 'De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data', *Cell Stem Cell*, 19: 266-77.

Hashimshony, T., N. Senderovich, G. Avital, A. Klochendler, Y. de Leeuw, L. Anavy, D. Gennert, S. Li, K. J. Livak, O. Rozenblatt-Rosen, Y. Dor, A. Regev, and I. Yanai. 2016. 'CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq', *Genome Biol*, 17: 77.

Hutchins, J. R., Y. Toyoda, B. Hegemann, I. Poser, J. K. Heriche, M. M. Sykora, M. Augsburg, O. Hudecz, B. A. Buschhorn, J. Bulkescher, C. Conrad, D. Comartin, A. Schleiffer, M. Sarov, A. Pozniakovsky, M. M. Slabicki, S. Schloissnig, I. Steinmacher, M. Leuschner, A. Ssykor, S. Lawo, L. Pelletier, H. Stark, K. Nasmyth, J. Ellenberg, R. Durbin, F. Buchholz, K. Mechtler, A. A. Hyman, and J. M. Peters. 2010. 'Systematic analysis of human protein complexes identifies chromosome segregation proteins', *Science*, 328: 593-9.

Jesty, S. A., M. A. Steffey, F. K. Lee, M. Breitbach, M. Hesse, S. Reining, J. C. Lee, R. M. Doran, A. Y. Nikitin, B. K. Fleischmann, and M. I. Kotlikoff. 2012. 'c-kit+ precursors support postinfarction myogenesis in the neonatal, but not adult, heart', *Proc Natl Acad Sci U S A*, 109: 13380-5.

Kasahara, H., S. Bartunkova, M. Schinke, M. Tanaka, and S. Izumo. 1998. 'Cardiac and extracardiac expression of Csx/Nkx2.5 homeodomain protein', *Circ Res*, 82: 936-46.

Kikuchi, K., J. E. Holdway, A. A. Werdich, R. M. Anderson, Y. Fang, G. F. Egnaczyk, T. Evans, C. A. Macrae, D. Y. Stainier, and K. D. Poss. 2010. 'Primary contribution to zebrafish heart regeneration by gata4(+) cardiomyocytes', *Nature*, 464: 601-5.

Kim, J. E., K. Nakashima, and B. de Crombrugghe. 2004. 'Transgenic mice expressing a ligand-inducible cre recombinase in osteoblasts and odontoblasts: a new tool to examine physiology and disease of postnatal bone and tooth', *Am J Pathol*, 165: 1875-82.

Kimura, W., F. Xiao, D. C. Canseco, S. Muralidhar, S. Thet, H. M. Zhang, Y. Abderrahman, R. Chen, J. A. Garcia, J. M. Shelton, J. A. Richardson, A. M. Ashour, A. Asaithamby, H. Liang, C. Xing, Z. Lu, C. C. Zhang, and H. A. Sadek. 2015. 'Hypoxia fate mapping identifies cycling cardiomyocytes in the adult heart', *Nature*, 523: 226-30.

Kretzschmar, K., and F. M. Watt. 2012. 'Lineage tracing', *Cell*, 148: 33-45.

Li, G. H., Y. Shi, Y. Chen, M. Sun, S. Sader, Y. Maekawa, S. Arab, F. Dawood, M. Chen, G. De Couto, Y. Liu, M. Fukuoka, S. Yang, M. Da Shi, L. A. Kirshenbaum, C. A. McCulloch, and P. Liu. 2009. 'Gelsolin regulates cardiac remodeling after myocardial infarction through DNase I-mediated apoptosis', *Circ Res*, 104: 896-904.

Love, M. I., W. Huber, and S. Anders. 2014. 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol*, 15: 550.

Madisen, L., T. A. Zwingman, S. M. Sunkin, S. W. Oh, H. A. Zariwala, H. Gu, L. L. Ng, R. D. Palmiter, M. J. Hawrylycz, A. R. Jones, E. S. Lein, and H. Zeng. 2010. 'A robust and high-throughput Cre reporting and characterization system for the whole mouse brain', *Nat Neurosci*, 13: 133-40.

Maruyama, S., K. Nakamura, K. N. Papanicolaou, S. Sano, I. Shimizu, Y. Asaumi, M. J. van den Hoff, N. Ouchi, F. A. Recchia, and K. Walsh. 2016. 'Follistatin-like 1 promotes cardiac fibroblast activation and protects the heart from rupture', *EMBO Mol Med*, 8: 949-66.

Mohamed, T. M. A., Y. S. Ang, E. Radzinsky, P. Zhou, Y. Huang, A. Elfenbein, A. Foley, S. Magnitsky, and D. Srivastava. 2018. 'Regulation of Cell Cycle to Stimulate Adult Cardiomyocyte Proliferation and Cardiac Regeneration', *Cell*, 173: 104-16 e12.

Muraro, M. J., G. Dharmadhikari, D. Grun, N. Groen, T. Dielen, E. Jansen, L. van Gurp, M. A. Engelse, F. Carlotti, E. J. de Koning, and A. van Oudenaarden. 2016. 'A Single-Cell Transcriptome Atlas of the Human Pancreas', *Cell Syst*, 3: 385-94 e3.

Nakada, Y., D. C. Canseco, S. Thet, S. Abdisalaam, A. Asaithamby, C. X. Santos, A. M. Shah, H. Zhang, J. E. Faber, M. T. Kinter, L. I. Szweda, C. Xing, Z. Hu, R. J. Deberardinis, G. Schiattarella, J. A. Hill, O. Oz, Z. Lu, C. C. Zhang, W. Kimura, and H. A. Sadek. 2017. 'Hypoxia induces heart regeneration in adult mice', *Nature*, 541: 222-27.

Naqvi, N., M. Li, J. W. Calvert, T. Tejada, J. P. Lambert, J. Wu, S. H. Kesteven, S. R. Holman, T. Matsuda, J. D. Lovelock, W. W. Howard, S. E. Iismaa, A. Y. Chan, B. H. Crawford, M. B. Wagner, D. I. Martin, D. J. Lefer, R. M. Graham, and A. Husain. 2014. 'A proliferative burst during preadolescence establishes the final cardiomyocyte number', *Cell*, 157: 795-807.

Pfister, O., A. Oikonomopoulos, K. I. Sereti, R. L. Sohn, D. Cullen, G. C. Fine, F. Mouquet, K. Westerman, and R. Liao. 2008. 'Role of the ATP-binding cassette transporter Abcg2 in the phenotype and function of cardiac side population cells', *Circ Res*, 103: 825-35.

Pinto, A. R., A. Ilinykh, M. J. Ivey, J. T. Kuwabara, M. L. D'Antoni, R. Debuque, A. Chandran, L. Wang, K. Arora, N. A. Rosenthal, and M. D. Tallquist. 2016. 'Revisiting Cardiac Cellular Composition', *Circ Res*, 118: 400-9.

Ranjbarvaziri, S., S. Park, N. B. Nguyen, W. B. Gilmore, P. Zhao, and R. Ardehali. 2017. 'Generation of Nkx2-5/CreER transgenic mice for inducible Cre expression in developing hearts', *Genesis*.

Scialdone, A., K. N. Natarajan, L. R. Saraiva, V. Proserpio, S. A. Teichmann, O. Stegle, J. C. Marioni, and F. Buettner. 2015. 'Computational assignment of cell-cycle stage from single-cell transcriptome data', *Methods*, 85: 54-61.

Senyo, S. E., M. L. Steinhauser, C. L. Pizzimenti, V. K. Yang, L. Cai, M. Wang, T. D. Wu, J. L. Guerquin-Kern, C. P. Lechene, and R. T. Lee. 2013. 'Mammalian heart renewal by pre-existing cardiomyocytes', *Nature*, 493: 433-6.

Sereti, K. I., N. B. Nguyen, P. Kamran, P. Zhao, S. Ranjbarvaziri, S. Park, S. Sabri, J. L. Engel, K. Sung, R. P. Kulkarni, Y. Ding, T. K. Hsiai, K. Plath, J. Ernst, D. Sahoo, H. K. A. Mikkola, M. L. Iruela-Arispe, and R. Ardehali. 2018. 'Analysis of cardiomyocyte clonal expansion during mouse heart development and injury', *Nat Commun*, 9: 754.

Skelton, R. J., M. Costa, D. J. Anderson, F. Bruveris, B. W. Finnin, K. Koutsis, D. Arasaratnam, A. J. White, A. Rafii, E. S. Ng, A. G. Elefanty, E. G. Stanley, C. W. Pouton, J. M. Haynes, R. Ardehali, R. P. Davis, C. L. Mummery, and D. A. Elliott. 2014. 'SIRPA, VCAM1 and CD34 identify discrete lineages during early human cardiovascular development', *Stem Cell Res*, 13: 172-9.

Smart, N., S. Bollini, K. N. Dube, J. M. Vieira, B. Zhou, S. Davidson, D. Yellon, J. Riegler, A. N. Price, M. F. Lythgoe, W. T. Pu, and P. R. Riley. 2011. 'De novo cardiomyocytes from within the activated adult heart after injury', *Nature*, 474: 640-4.

Soonpaa, M. H., D. C. Zebrowski, C. Platt, A. Rosenzweig, F. B. Engel, and L. J. Field. 2015. 'Cardiomyocyte Cell-Cycle Activity during Preadolescence', *Cell*, 163: 781-2.

Sylva, M., V. S. Li, A. A. Buffing, J. H. van Es, M. van den Born, S. van der Velden, Q. Gunst, J. H. Koolstra, A. F. Moorman, H. Clevers, and M. J. van den Hoff. 2011. 'The BMP antagonist follistatin-like 1 is required for skeletal and lung organogenesis', *PLoS One*, 6: e22616.

Tallquist, M. D., and J. D. Molkentin. 2017. 'Redefining the identity of cardiac fibroblasts', *Nat Rev Cardiol*.

Uchida, S., P. De Gaspari, S. Kostin, K. Jenniches, A. Kilic, Y. Izumiya, I. Shiojima, K. Grosse Kreymborg, H. Renz, K. Walsh, and T. Braun. 2013. 'Sca1-derived cells are a source of myocardial renewal in the murine adult heart', *Stem Cell Reports*, 1: 397-410.

Uygur, A., and R. T. Lee. 2016. 'Mechanisms of Cardiac Regeneration', *Dev Cell*, 36: 362-74.

van Berlo, J. H., O. Kanisicak, M. Maillet, R. J. Vagnozzi, J. Karch, S. C. Lin, R. C. Middleton, E. Marban, and J. D. Molkentin. 2014. 'c-kit+ cells minimally contribute cardiomyocytes to the heart', *Nature*, 509: 337-41.

van Berlo, J. H., O. Kanisicak, M. Maillet, R. J. Vagnozzi, J. Karch, S. J. Lin, R. C. Middleton, E. Marban, and J. D. Molkentin. 2018. 'van Berlo et al. reply', *Nature*, 555: E18.

van Berlo, J. H., and J. D. Molkentin. 2014. 'An emerging consensus on cardiac regeneration', *Nat Med*, 20: 1386-93.

Vicinanza, C., I. Aquila, E. Cianflone, M. Scalise, F. Marino, T. Mancuso, F. Fumagalli, E. D. Giovannone, F. Cristiano, E. Iaccino, P. Marotta, A. Torella, R. Latini, V. Agosti, P. Veltri, K. Urbanek, A. M. Isidori, D. Saur, C. Indolfi, B. Nadal-Ginard, and D. Torella. 2018. 'Kit(cre) knock-in mice fail to fate-map cardiac stem cells', *Nature*, 555: E1-E5.

Vicinanza, C., I. Aquila, M. Scalise, F. Cristiano, F. Marino, E. Cianflone, T. Mancuso, P. Marotta, W. Sacco, F. C. Lewis, L. Couch, V. Shone, G. Gritti, A. Torella, A. J. Smith, C. M. Terracciano, D. Britti, P. Veltri, C. Indolfi, B. Nadal-Ginard, G. M. Ellison-Hughes, and D. Torella. 2017. 'Adult cardiac stem cells are multipotent and robustly myogenic: c-kit expression is necessary but not sufficient for their identification', *Cell Death Differ*.

Wei, K., V. Serpooshan, C. Hurtado, M. Diez-Cunado, M. Zhao, S. Maruyama, W. Zhu, G. Fajardo, M. Noseda, K. Nakamura, X. Tian, Q. Liu, A. Wang, Y. Matsuura, P. Bushway, W. Cai, A. Savchenko, M. Mahmoudi, M. D. Schneider, M. J. van den Hoff, M. J. Butte, P. C. Yang, K. Walsh, B. Zhou, D. Bernstein, M. Mercola, and P. Ruiz-Lozano. 2015. 'Epicardial FSTL1 reconstitution regenerates the adult mammalian heart', *Nature*, 525: 479-85.

Whitfield, M. L., L. K. George, G. D. Grant, and C. M. Perou. 2006. 'Common markers of proliferation', *Nat Rev Cancer*, 6: 99-106.

Yellamilli, A., and J. H. van Berlo. 2016. 'The Role of Cardiac Side Population Cells in Cardiac Regeneration', *Front Cell Dev Biol*, 4: 102.

Zeisberg, E. M., and R. Kalluri. 2010. 'Origins of cardiac fibroblasts', *Circ Res*, 107: 1304-12.

SUPPLEMENTARY INFORMATION
Surgical Procedures
For sham, myocardial infarction (MI) and ischaemia/reperfusion (I/R) surgeries, 8-week old mice were anesthetised with mixture of ketamine and xylazine by intra-peritoneal (IP) injection. The fur was removed from the ventral surface of the neck and thorax with Nair hair removal cream (Church & Dwight). Mice were then artificially ventilated via a tracheal tube connected to a ventilator. After cleaning the surgical site with iodine and 70% ethanol, the skin was incised left of the midline with allowing access to the third intercostal space. Next, the pectoral muscles were retracted, the intercostal muscles were cut caudal to the third rib and a rib spreader was applied to allow access to the heart. The pericardium was incised longitudinally and the left anterior descending coronary artery (LAD) was identified.

For the sham control surgery, the rib cage was then closed with 5-0 silk suture without placing a ligature around the LAD and the skin was closed with tissue adhesive.

For the MI injury, a 7-0 silk suture was placed beneath the LAD to permanently occlude the artery. The rib cage was closed with 5-0 silk suture and the skin was closed with tissue adhesive.

For the I/R injury, a 2–3 mm piece of PE 10 tubing was placed over the LAD and the ligature secured around the LAD and PE tubing. The rib cage was closed with 5-0 silk suture and the skin was closed with tissue adhesive. Following 1-hour ischaemic period, the PE tubing was removed and the ligature was cut to allow for reperfusion via the LAD. Again, the rib cage was closed with 5-0 silk suture and the skin was closed with tissue adhesive.

After these surgical procedures, the mice were disconnected from the ventilator, the tracheal tube was removed and the animal was placed unrestrained on a nose cone with 100% oxygen in a warm recovery cage until fully ambulatory, at which point the oxygen was turned off. To alleviate pain or distress Buprenorphine was injected subcutaneously as analgesic (0.05-0.1 mg·kg^{-1} bodyweight), given once at completion of surgery. Mice were killed at time points indicated below or in the main text.

Genetic Lineage-Tracing Experiments
For genetic lineage tracing of *Mki67*-expressing cells in neonatal hearts, 1-week old *Mki67*[IRES-CreERT2] × LSL-tdTomato mice were treated once with 150 µg tamoxifen (Sigma-Aldrich) dissolved in sunflower oil via the IP route. Mice were killed 1 day and 7 days after tamoxifen injection and tissue was collected for histological analysis.

For genetic lineage tracing of *Mki67*-expressing cells in adult homoeostatic hearts, 8-week old *Mki67*[IRES-CreERT2] × LSL-tdTomato mice were treated either once or five consecutive times with 5 mg tamoxifen dissolved in sunflower oil via IP injection. Mice were killed (1 day after tamoxifen injection or at 1.3–1.5 years of age) and tissue was collected for histological analysis.

For genetic lineage tracing of *Mki67*-expressing cells in upon myocardial infarction surgery, 8-week old *Mki67*[IRES-CreERT2] × LSL-tdTomato mice that had undergone sham or MI surgery were IP injected with 5 mg tamoxifen dissolved in sunflower oil every other day for

3 doses in total. Mice were killed 7 and 28 days following surgery and tissue was collected for histological analysis.

For genetic lineage tracing of *Fstl1*-expressing cells in upon myocardial infarction surgery, 7-week old *Fstl1*[eGFP-IRES-CreERT2] × LSL-tdTomato were IP injected once with 5 mg tamoxifen dissolved in sunflower oil. One week later, MI surgery was performed. Mice were killed 2 weeks following MI surgery and tissue was collected for histological analysis.

Cardiac Cell Isolation

Mouse hearts were dissected, the atria were removed and, when suitable, hearts were divided up into remote and apex (*i.e.* for MI or I/R surgery, borderzone and infarcted zone). Whole ventricles were prepped from dissected hearts of neonatal mice. Dissected tissue was minced into small pieces (<1 mm) and incubated for 15 minutes at 37 °C (water bath) in 0.5 mM ethylenediaminetetraacetic acid (EDTA) in phosphate buffered saline without Ca^{2+} and Mg^{2+} (PBSO). Incubated pieces were disrupted using pre-wetted glass Pasteur pipettes by repeatedly pipetting up and down (>10 times). Pieces were centrifuged at $100 \times g$ for 5 minutes at 4 °C and the supernatant was carefully aspirated. Tissue pieces were washed in PBS and again centrifuged at $100 \times g$ for 5 minutes at 4 °C. After supernatant removal, tissue pieces were resuspended in dissociation buffer: Advanced DMEM/F12 containing collagenase II (1 mg·mL⁻¹; Thermo Scientific), DNase I (10 U·mL⁻¹; Sigma-Aldrich) and hyaluronidase (10 µg·mL⁻¹; Sigma-Aldrich). Samples were incubated for 20-30 minutes at 37 °C on a rocker and then further disrupted using pre-wetted glass Pasteur pipettes by repeatedly pipetting up and down (>10 times). To inactivate enzymatic activity, foetal calf serum (FCS) was added to the cell suspension, which was then centrifuged at $300 \times g$ for 5 minutes at 4 °C. After supernatant removal, the cell pellet was incubated in 1× red blood cell (RBC) lysis solution (freshly prepared from a 10× stock solution containing 1.55M NH_4Cl, 100 mM $KHCO_3$ and 0.1 mM EDTA dissolved in MilliQ water) for 5 minutes at room temperature. Upon RBC lysis, cells were centrifuged at $300 \times g$ for 5 minutes at 4 °C and washed in PBS, followed by another centrifugation at $300 \times g$ for 5 minutes at 4 °C. Cells were stained with MitoTracker Deep Red (Cell Signaling) for 10 minutes at 37 °C on a shaker. Cells were resuspended in cold flow cytometry buffer (PBS, 5% FCS, 2 mM EDTA) and filtered through a 70-µm cell strainer (Greiner) before proceeding with flow sorting. Cardiac cell isolation from some wildtype (*Mki67*[wt/wt]) mice was performed as recently described (1).

Messenger RNA Sequencing

Samples were lysed and RNA from each bulk sort sample or single cell was barcoded and processed using the CEL-Seq2 technique (2-4). Briefly, for bulk sequencing, samples sorted into TRIzol were extracted following the manufacturer's instruction, except, total RNA was precipitated with 2 µg GlycoBlue (Ambion) overnight at −80 °C. RNA pellets were dissolved in reverse transcription reaction mix (Invitrogen) containing UMI barcode primers and dNTPs (Promega) and incubated at 70 °C for 2 minutes. For single-cell mRNA-sequencing, cells were directly lysed at 65 °C for 5 minutes, followed by first and second strand synthesis (Invitrogen).

Sequencing samples were pooled into single library prior to *in vitro* transcription (Ambion). Amplified RNA was used as template to generate complementary DNA (cDNA) libraries using Illumina TruSeq primers. Libraries were sequenced on an Illumina NextSeq500 using 75-bp pair-end sequencing. Bioinformatics analysis of sequenced libraries was performed as described below.

Histology

Tissues were fixed in 4% paraformaldehyde (PFA) in PBS at room temperature for 2 hours. After fixation, the hearts were incubated in cryoprotective solution (30% sucrose in PBS) at 4 °C overnight, then frozen blocks were prepared in tissue freezing medium (Leica) and stored at −80 °C. Ten μm thick cryosections were cut and mounted on cover glass. Dried sections were fixed using 4% paraformaldehyde for 15 minutes, permeabilised with 0.5% Triton X-100 in PBS for 5 min and blocked for 2 hours in PBS containing 10% FCS, 3% bovine serum albumin and 0.5% fish skin gelatin (Sigma-Aldrich). Next, sections were incubated overnight with primary antibody in blocking solution, washed and incubated with secondary antibody for 2 hours at room temperature. Finally, sections were washed and mounted in ProLong Gold Antifade Mounting Medium with DAPI (Life Technologies). For paraffin blocks, hearts were fixed using 4% PFA, dehydrated using ethanol and embedded in paraffin. Five μm-sections were processed using standard methods and antigen retrieval was performed using pH 6 citrate buffer. Dewaxed sections were incubated with primary antibody overnight at 4 °C and then incubated for 1 hour at room temperature with BrightVision poly-HRT anti-rabbit (Immunologic). Immunoreactivity was visualized using 3,3'-diaminobenzidine (DAB).

In Situ Hybridisation

In situ hybridisation (ISH) was performed on 8 μm paraffin sections of mouse hearts, as described (5). Sections were deparaffinised, rehydrated in a series of alcohol and incubated for 15 minutes in 20 μg/mL proteinase K in PBS for 15 minutes at 37 °C. Proteinase K was inactivated using 0.2% glycine in PBS, after which the sections were washed in PBS, fixed in 4% PFA and 0.2% glutaraldehyde in PBS. Sections were incubated overnight with the following probes in hybridization mix: *Col1a1* (0.5 ng·μL^{-1}) and *Col3a1* (0.5 ng·μL^{-1}) After hybridisation, the sections were washed in 2× SSC (150mM NaCl, 15mM Na-citrate, pH 7), 50% formamide in 2× SSC pH 7 and TNT (150 mM NaCl, 100 mM Tris-HCl, 0.05% Tween-20, pH 7.5). Blocking was done in MABT (150 mM NaCl, 100 mM maleic acid, 0.1% Tween-20, pH 7.4) plus 2% Block Reagent (Roche 1096176), followed by 2 hours incubation with anti-DIG-alkaline in MABT-Block 2% (1:1500), washed in TNT and NTM (100 mM NaCl, 100 mM Tris-HCl, 50 mM MgCl$_2$, pH9.2), and subsequently followed by incubation with NBT/BCIP (1:50). After colour development the sections were rinsed with double distilled water, dehydrated in a graded series of ethanol, washed in xylene and mounted in Entellan.

Bioinformatics Analysis

DNA library sequencing, mapping to the mouse reference genome and quantification of transcript abundance were performed as described elsewhere (4). Bulk sequencing libraries were analysed using the DESeq2 package (6). Cell clusters and gene expression levels across clusters were generated using the RaceID2 algorithm (3) from single-cell libraries. For initial cell type analysis, ERCC92 spike-ins and mitochondrial genes and cell transcriptomes with less than 1,000 uniquely expressed transcripts were removed from the single-cell dataset. Single-cell sequencing libraries were normalised by downsampling to a minimum number of 1,000 transcripts. After first running RaceID2, 2,144 cells making the 1,000-transcript threshold separated into 14 cell clusters. Subsequently, cell clusters enriched for *Kcnq1ot1* (cluster 5) and *Rn45s* (cluster 13), markers of necrotic cells with low quality RNA, were removed followed by running RaceID2 again to generate the cell clusters presented in this paper (7, 8). For refined analysis of cardiac cell subpopulations, cells assigned to the five main cardiac cell types were extracted and, subsequently, cells with more than 2,000 transcript counts assigned to the lineages of cardiomyocytes and cardiac fibroblasts were used for further analysis. To increase robustness, clusters containing less than 10 cells were removed from further analysis (8): this applied to the initial subclustering of cardiomyocytes, where clusters 5, 7 and 9 (of a total of 11 clusters) were not considered for the subsequent differentially gene-expression analysis. After removal of these clusters, RaceID2 was run again, which resulted in 6 clusters. All bioinformatics analysis was performed using R version 3.4.0 (R Foundation, https://www.r-project.org) and RStudio version 1.0.143 (https://www.rstudio.com).

REFERENCES FOR SI REFERENCE CITATIONS

1. Gladka MM, *et al.* (2018) Single-Cell Sequencing of the Healthy and Diseased Heart Reveals Ckap4 as a New Modulator of Fibroblasts Activation. *Circulation*.

2. Hashimshony T, *et al.* (2016) CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol* 17:77.

3. Grün D, *et al.* (2016) De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data. *Cell Stem Cell* 19(2):266-277.

4. Muraro MJ, *et al.* (2016) A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Syst* 3(4):385-394 e383.

5. Sylva M, *et al.* (2011) The BMP antagonist follistatin-like 1 is required for skeletal and lung organogenesis. *PLoS One* 6(8):e22616.

6. Love MI, Huber W, & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.

7. Scheele CL, *et al.* (2017) Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature* 542(7641):313-317.

8. Grün D, *et al.* (2015) Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* 525(7568):251-255.

Other Supporting Information Files can be found online:
https://www.pnas.org/content/115/52/E12245.short

Dataset S1. Differentially expressed genes of the initial cardiac cell clustering.

Dataset S2. Differentially expressed genes of the cardiomyocyte subclustering.

Dataset S3. TOP 100 differentially expressed genes in neonatal cardiomyocytes, adult injured-associated cardiomyocytes and adult homeostatic cardiomyocytes.

**Dataset S4.** Differentially expressed genes of the cardiac fibroblast subclustering.
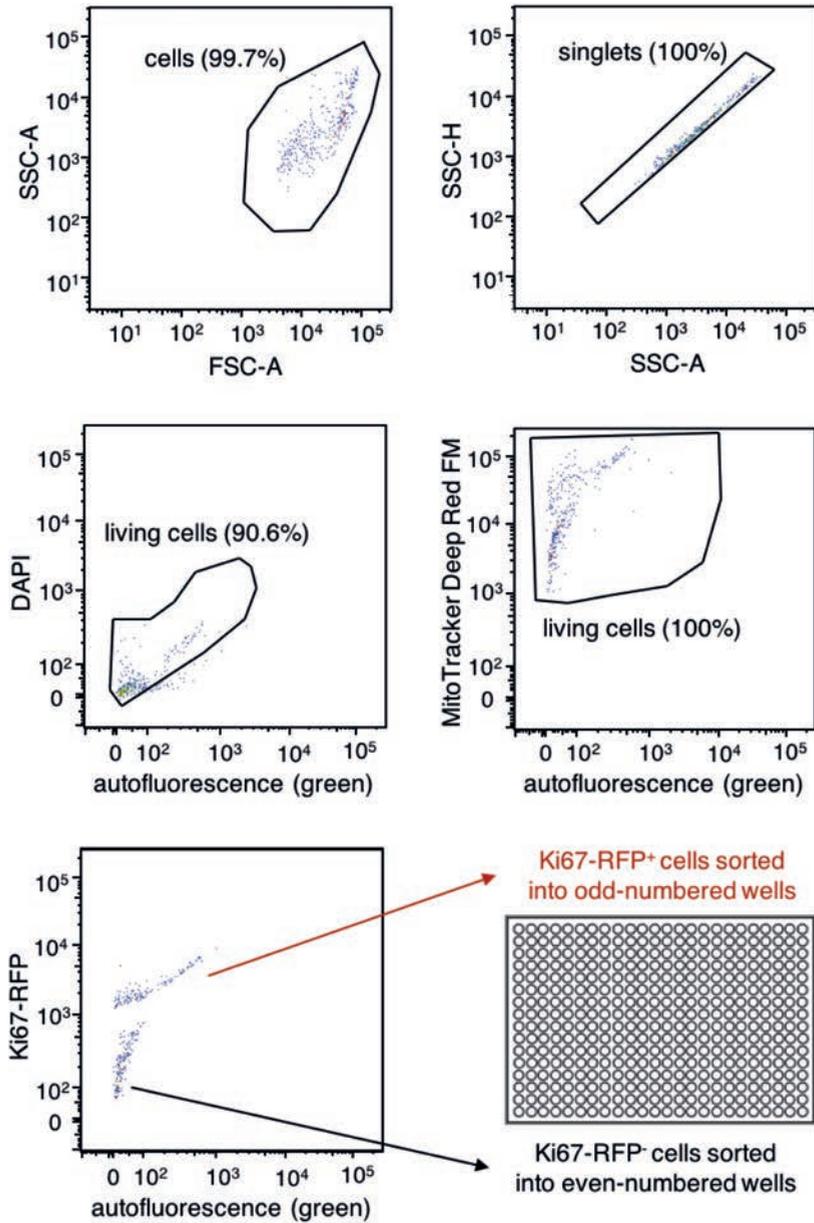
Figure S1. Gating and sorting strategy for the purification of Ki67-RFP+ cells. Representative dot plots showing the gating and sorting strategy (from left to right, from top to bottom). Area scaling, A; forward scatter, FSC; height scaling, H; side scatter, SSC.
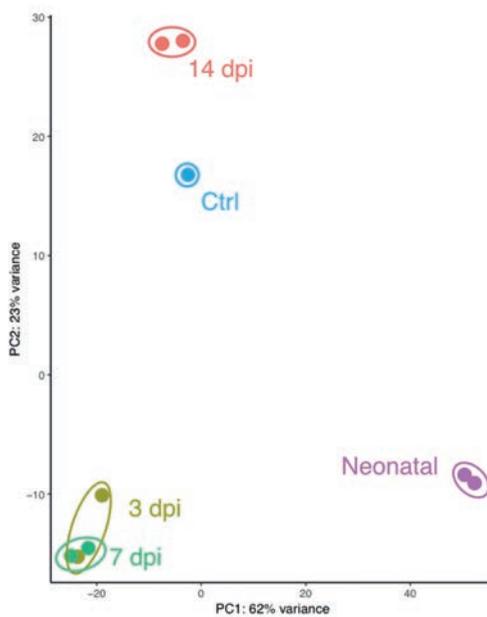
Figure S2. Principal component analysis (PCA) of bulk mRNA-sequencing of Ki67-RFP+ cells. Principal component analysis (PCA) plot of mRNA expression profiles of Ki67-RFP+ cells isolated at 7 days of age (Neonatal), 3, 7 and 14 days after myocardial infarction (dpi) and or 14 days after sham surgery (Ctrl).
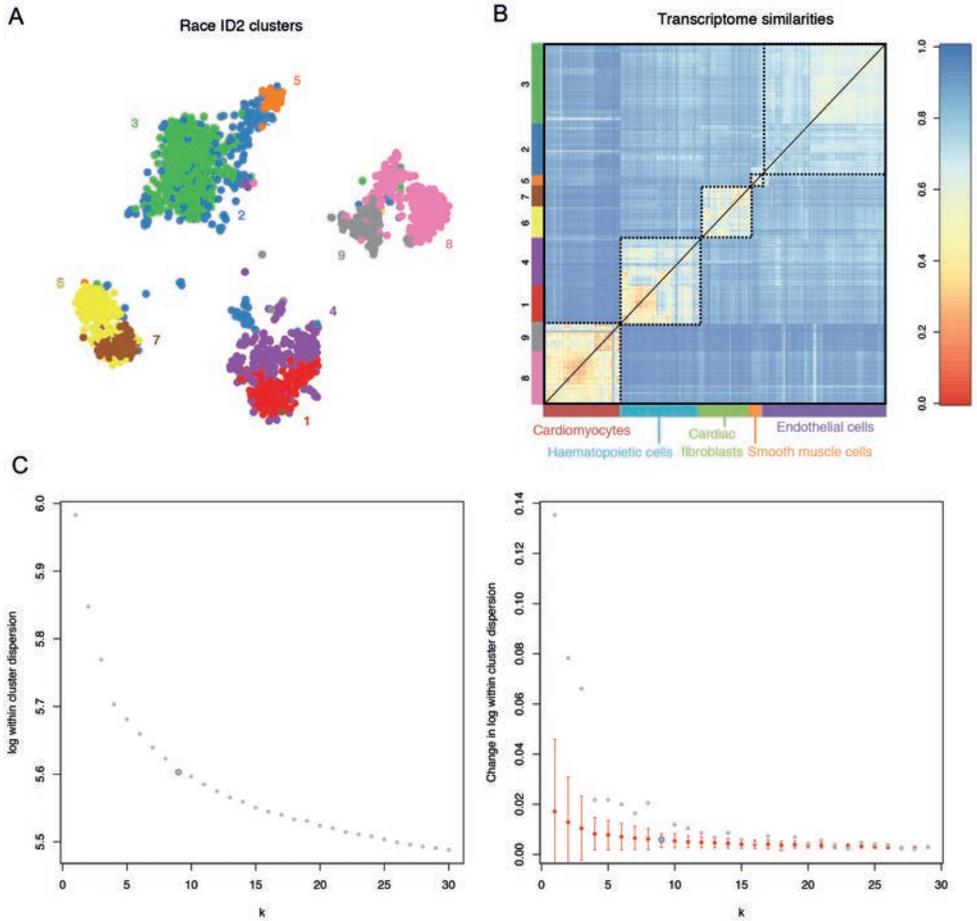
Figure S3. Cell clustering. (A) Clustering of cardiac cells and cell-to-cell distances visualized by t-distributed stochastic neighbour-embedding (t-SNE) map, highlighting 9 identified clusters. (B) Heatmap representation of the transcriptome similarities between 1,939 single cells from the 9 identified clusters. (C) k-medoids clustering. If the change of the within-cluster dispersion upon increasing the cluster number ($k_{i+1} = k_i + 1$) is within the error of the average change upon further increase (k_{i+2}, …, k_{max}), k_i is chosen as input. The average change across cluster numbers k_{i+2}, …, k_{max} and its error is computed from a linear regression. Within-cluster dispersion (see Figure below, left) and change of the within-cluster dispersion (see Figure below, right) as a function of the number of clusters. The average change at cluster number k upon further increase is shown in red. In both panels the selected cluster number is circled in blue.
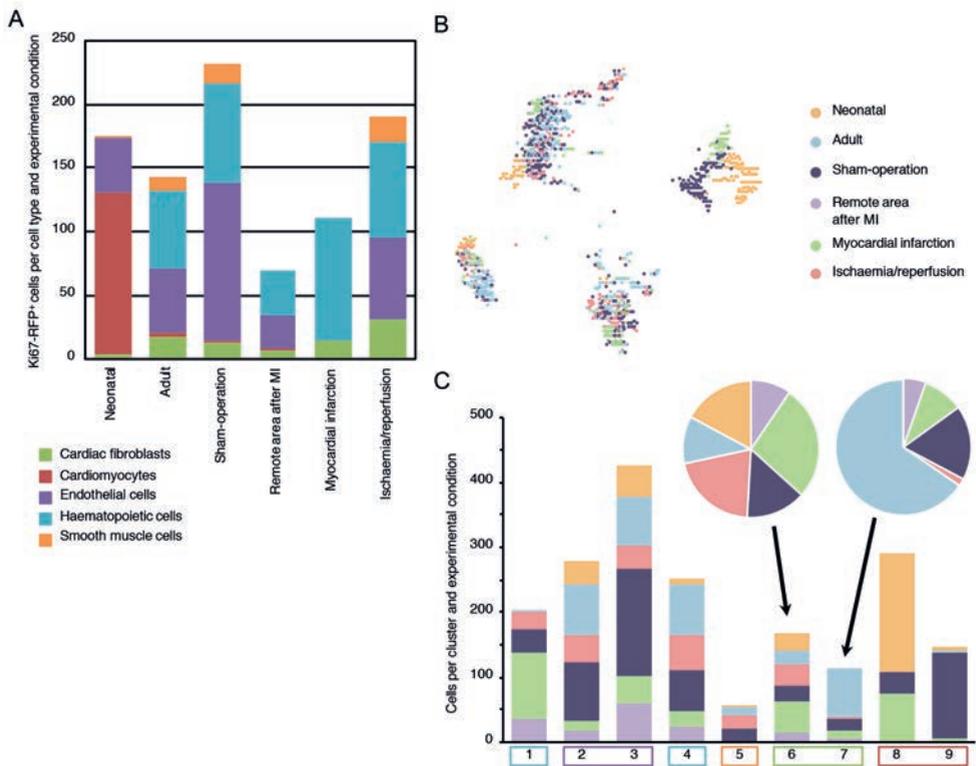
Figure S4. Expression patterns of cardiac lineage genes in sequenced cardiac single cells. (A) Ki67-RFP+ cells per experimental group assigned to cell types based on marker genes enriched in each major cell cluster. (B) t-SNE map representation of the experimental condition for each sequenced cell. (C) Quantification of the conditions of origin shown on B within each cell cluster (colour of the frame around the cell cluster number corresponds to the cell types shown in panel A).
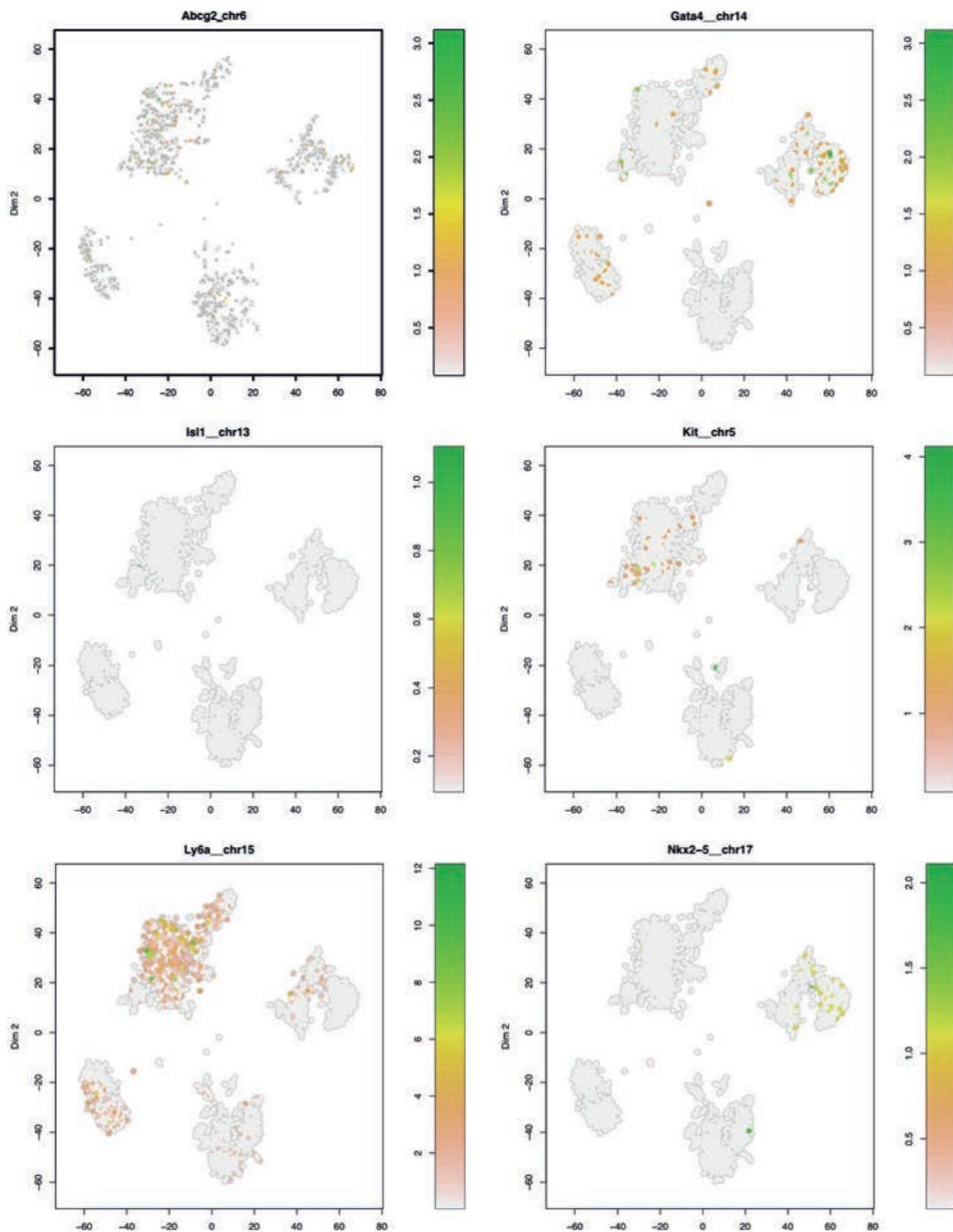
Figure S5. Expression (normalised transcript counts) of putative cardiac stem cell markers in the dataset. mRNA transcript counts of Abcg2, Gata4, Isl1, Kit (c-KIT), Ly6a (SCA-1) and Nkx2-5.
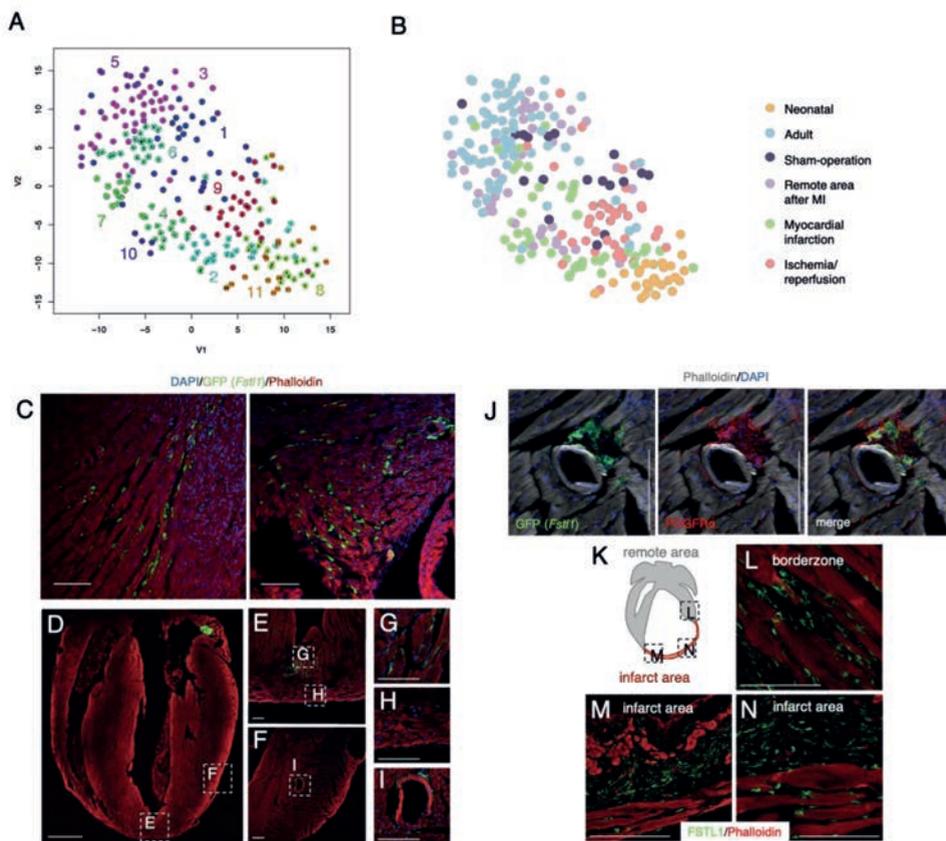
Figure S6. Fstl1 is highly expressed in fibroblasts from neonatal mice and after MI. (A) Clustering of cardiac fibroblasts and cell-to-cell distances visualised by t-distributed stochastic neighbour-embedding (t-SNE) map, highlighting 11 identified clusters. (B) Projection of the experimental conditions over the fibroblast sub-clusters. (C-J) eGFP+ cells (green) in hearts of neonatal (C) and homoeostatic (D-J) Fstl1eGFP-IRES-CreERT2 mice overlap with fibroblast marker PDGFRα (J, red) and no eGFP+ were found in the epicardium (H). Nuclei were counterstained with DAPI (blue) and cell borders visualized by Phalloidin (as indicated, either red or white). (K-N) Validation of FSTL1 expression in the borderzone and in the infarcted region of wildtype mice two after weeks after MI using a specific anti- FSTL1 antibody (green) and cell border visualised by Phalloidin (red). Scale bars: 100 m (C, E-J, K-N) and 1 mm (D).

A HUMAN TEAR GLAND ATLAS AND ORGANOIDS THAT CAN CRY

3
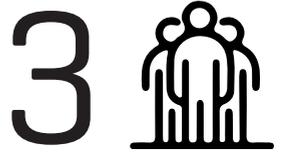
Yorick Post[1], Marie Bannier-Hélaouët[1,2], Marc T. Bustos[1], Yotam E. Bar-Ephraim[1,2], Jelte van der Vaart[1,2], Harry Begthel[1], Jeroen Korving[1], Rachel Kalmann[3], Saskia M. Imhoff[3] and Hans Clevers[1,2,4]

[1] Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences, University Medical Center Utrecht, Utrecht, the Netherlands.
[2] Oncode Institute, Utrecht, the Netherlands.
[3] Department of Ophthalmology, University Medical Center Utrecht, Utrecht, The Netherlands.
[4] Correspondence should be addressed to H.C. (email h.clevers@hubrecht.eu)

ABSTRACT

The lacrimal gland is essential for lubrication and protection of the eye. Disruption of lacrimal fluid production, composition or release results in dry eye, causing discomfort and damage to the eye surface. Here, we provide a single-cell RNA expression atlas of the human tear gland and unveil its cellular heterogeneity. We describe the establishment of a long-term 3D organoid culture system for mouse and human lacrimal gland. Organoids can be expanded for multiple months and recapitulate morphological and transcriptional features of the primary tissue. CRISPR/Cas9 mediated genome editing reveals the master control gene for eye development *Pax6* to be required for differentiation of adult lacrimal gland cells. A functional swelling assay of human organoids in response to neurotransmitters phenocopies the process of tear secretion. This study provides an experimental platform to study the (patho-)physiology of the tear gland.

Keywords

INTRODUCTION

Lacrimal glands are exocrine organs that secrete the aqueous layer of the tear film, one for each eye. Tears consists of water, mucins, electrolytes and antibacterial proteins. Different types of tears exist: basal tears are constitutively secreted to lubricate and moisten the eye. Reflex tears are produced by external stimuli to flush away irritants, while emotional tears are triggered by strong positive or negative emotions (Murube 2009). Tear fluid is secreted by the epithelial cells of the main lacrimal gland, which consists of acini and intralobular ducts (Hodges and Dartt 2003). Dysfunction in tear production or secretion potentially results in - or can be a consequence of - pathologies of the eye, such as Sjögren's syndrome (Brito-Zeron et al. 2016). The versatile function of tears is reflected in its complex chemical composition (de Souza, Godoy, and Mann 2006). Tear fluid contains many different proteins that are described to reflect normal and disease states (Perumal et al. 2016). The cellular heterogeneity of lacrimal gland epithelium, underlying tear components, have remained largely unexplored.

RESULTS

While cultures of primary tear gland cells have been described, they are either short-lived (Lin, Liu, and Yiu 2019), the growth medium not chemically defined (Tiwari et al. 2012) or derived from embryonic material (Hirayama et al. 2013). We have previously defined 3D culture conditions that allow long-term expansion of mouse and human organoids from multiple adult organs (Clevers 2016). To generate lacrimal gland organoids, we applied our adult stem cell-based organoid protocol from other organs (Drost et al. 2016; Hu et al. 2018; Sato et al. 2009; Schutgens et al. 2019) to mouse tissue isolates. Wild type murine lacrimal glands were dissociated into small fragments and embedded in basement membrane extract (BME) (**Figure 1A**). Supplying a 'generic' organoid medium containing R-spondin 3, Noggin, FGF10, TGF beta inhibitor A83-01 and cAMP/PKA activators PGE2 and FSK resulted in the most robust organoid outgrowth (**Figure S1A, B**). Addition of Nicotinamide slightly reduced initial outgrowth from primary tissue but is described to increase the lifespan of established organoids and therefore retained in the cocktail (Sato et al. 2011). Using this 'expansion medium', organoids expanded to a diameter of ~200 μm within the first ten days (**Figure 1B**). Thereafter, organoids could be passaged by mechanical disruption at a ratio of 1:4 every 7 days, for multiple months (>20 passages) without significant changes in growth kinetics or morphology (**Figure 1B**). To investigate the similarity between lacrimal gland organoids and the tissue of origin, we performed bulk mRNA sequencing of established organoid lines (n=3), lacrimal gland tissue (n=3) and the adjacent salivary gland tissue (n=2) as a reference (**Figure S1C**). To establish a lacrimal gland gene-expression signature, we compared both organoids and tissue with the salivary gland. We found 100 genes commonly upregulated (fold-change > 1.5, padj < 0.05) in the samples originating from the lacrimal gland compared to the salivary gland (**Figure 1C**). Overlapping differentially expressed genes included epithelial marker (*Krt6a*), ocular transcription factor (*Foxc2*, (Smith et al. 2000)), aquaporin (*Aqp4*), secreted tear proteins (*Ltf, Muc4*), a nonvolatile peptide described to direct sociosexual communication in rodent tear

fluid (*Esp6*, (Kimoto et al. 2007)) and the master regulator of eye development (*Pax6*, (Gehring 1996)) (**Figure 1D**). Among the commonly upregulated genes we did not find acinar cell markers *Sox10* and *Mist1* (Farmer et al. 2017), suggesting the organoids represent the ductal compartment of the lacrimal gland. Mapping protein expression using immunohistochemistry we confirmed nuclear expression of transcription factor Pax6 in all organoid cells and in epithelial cells of the tissue (**Figure 1E**). The *Pax6* gene is expressed in different tissues of the eye, including the lacrimal gland epithelium (Makarenkova et al. 2000). In both vertebrates and invertebrates, Pax6 can direct the formation of ectopic eyes (Chow et al. 1999; Halder, Callaerts, and Gehring 1995). We detected a polarized apical expression pattern of aquaporin in the organoids, as seen in the ducts and acini *in vivo* (**Figure 1E**). Defects in apical distribution of AQP5 have been implicated with decreased lacrimation and dry eye in patients with Sjogren's syndrome (Tsubota et al. 2001). Lactotransferrin (Ltf), described to have antimicrobial activity in various secretory fluids (Jenssen and Hancock 2009), accumulated in the lumen of the organoids. In the tissue we
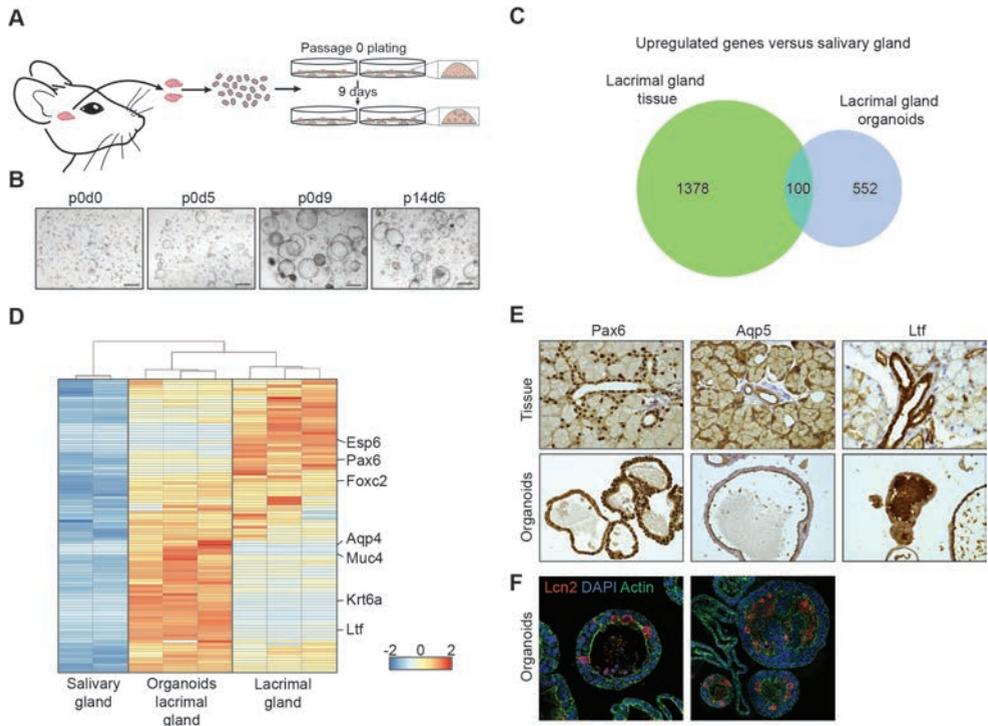

Figure 1. Establishment of organoid culture conditions for murine lacrimal gland. (A) Schematic representation of deriving organoids from murine lacrimal glands. (B) Organoid outgrowth in BME from plating primary tissue to late-passage. (C) Venn diagram showing overlap of 100 significantly upregulated genes (fc ≥ 1.5, p ≤ 0.05) lacrimal gland tissue and lacrimal gland organoids compared to salivary gland tissue. (D) Clustered heatmap of 100 upregulated genes from c. (E) Immunohistochemistry for lacrimal gland genes identified in bulk RNA sequencing (Pax6, Aqp5, Ltf) in tissue and organoids. (F) Immunofluorescent staining for Lcn2 (red) in lacrimal gland organoids.

found Ltf to be highly enriched in the ducts of the lacrimal gland, in contrast to other reports in human (Flanagan and Willcox 2009) (**Figure 1E**). A tear lactotransferrin assay determining protein concentration serves as a test for lacrimal gland function in both healthy and diseased individuals (Janssen and van Bijsterveld 1983). Tear component lipocalin 2 (Lcn2), an iron-sequestering protein in the antibacterial innate immune response, accumulated in luminal cells of the organoids (**Figure 1F**). The detectable concentration of lipocalin in tears is decreased in patients with Sjogren's syndrome (Caffery et al. 2008).

We noticed that larger organoids tended to differentiate into collapsed thicker-walled structures when fresh medium was not added frequently enough or cells were not split in time. This phenotype could be reproduced on a broader scale when withdrawing most growth factors and small molecules (R-spondin, Noggin, A83-01, PGE2, Nicotinamide and N-acetylcysteine) from the expansion medium (**Figure 2A**). Organoids became less proliferative, denser and remained viable for several days. Based on these morphological features we defined this as our 'differentiation medium'. Direct comparison of organoids in expansion medium (n=3) and organoids in differentiation medium (n=3) for 5 days using bulk mRNA sequencing revealed a transcriptional separation between the two conditions (**Figure S2A and B**). Upon differentiation, organoids decreased the expression of Wnt pathway components *Axin2*, *Fzd10* and marker of epithelial cells with stem cell potential *Tnfrsf19*, encoding Troy (Stange et al. 2013) (**Figure 2B**). Differentiated organoids upregulated the Wnt inhibitor *Dkk3*, luminal water channel *Aqp5*, marker of myoepithelium *Acta2* (**Figure 2B**), but also Keratin 20 (*Krt20*), established as a marker of mature cells in the gastric and intestinal epithelium (Chan et al. 2009) (**Figure 2B and S2B**). Differentiation did not increase the expression of acinar cell markers but rather induced a signature of ductal cell maturation and the myoepithelium lineage. This is in line with a lineage tracing study reporting Krt5+ basal cells giving rise to basal and luminal duct cells as well as myoepithelial cells but not acini (Farmer et al. 2017).

Transcription factor Pax6 is indispensable for lacrimal gland development in mouse (Makarenkova et al. 2000) and its overexpression crucial in the generation of lacrimal gland epithelium-like cells from human pluripotent stem cells (Hirayama et al. 2017). To decipher its role in adult lacrimal gland cells we utilized CRISPR/Cas9 technology for targeted deletion of *Pax6* in our organoids (**Figure 2C**). Clonally expanded knockout lines had undetectable levels of Pax6 protein by immunohistochemistry and western blot (**Figure 2D and E**). In expansion medium, Pax6-knockout (Pax6-KO) lines were indistinguishable from their wild type counterparts in culture and maintained the capacity for repeated passaging (**Figure 2F**). Loss of *Pax6* in adult lacrimal gland-derived organoids did not affect their viability, growth potential or cystic morphology (**Figure 2F**). Pax6-KO organoids decreased proliferation around 7 days after splitting but did not spontaneously differentiate. We next tested maturation capacity of Pax6-KO lines in response to differentiation medium. While wild type lines strongly reacted (as described above), Pax6-KO lines stopped proliferating but retained an undifferentiated cystic phenotype upon exposure to differentiation medium (**Figure 2G**). In line with the morphological phenotype, Pax6-KO lines did not upregulated *Krt20* mRNA in differentiation medium
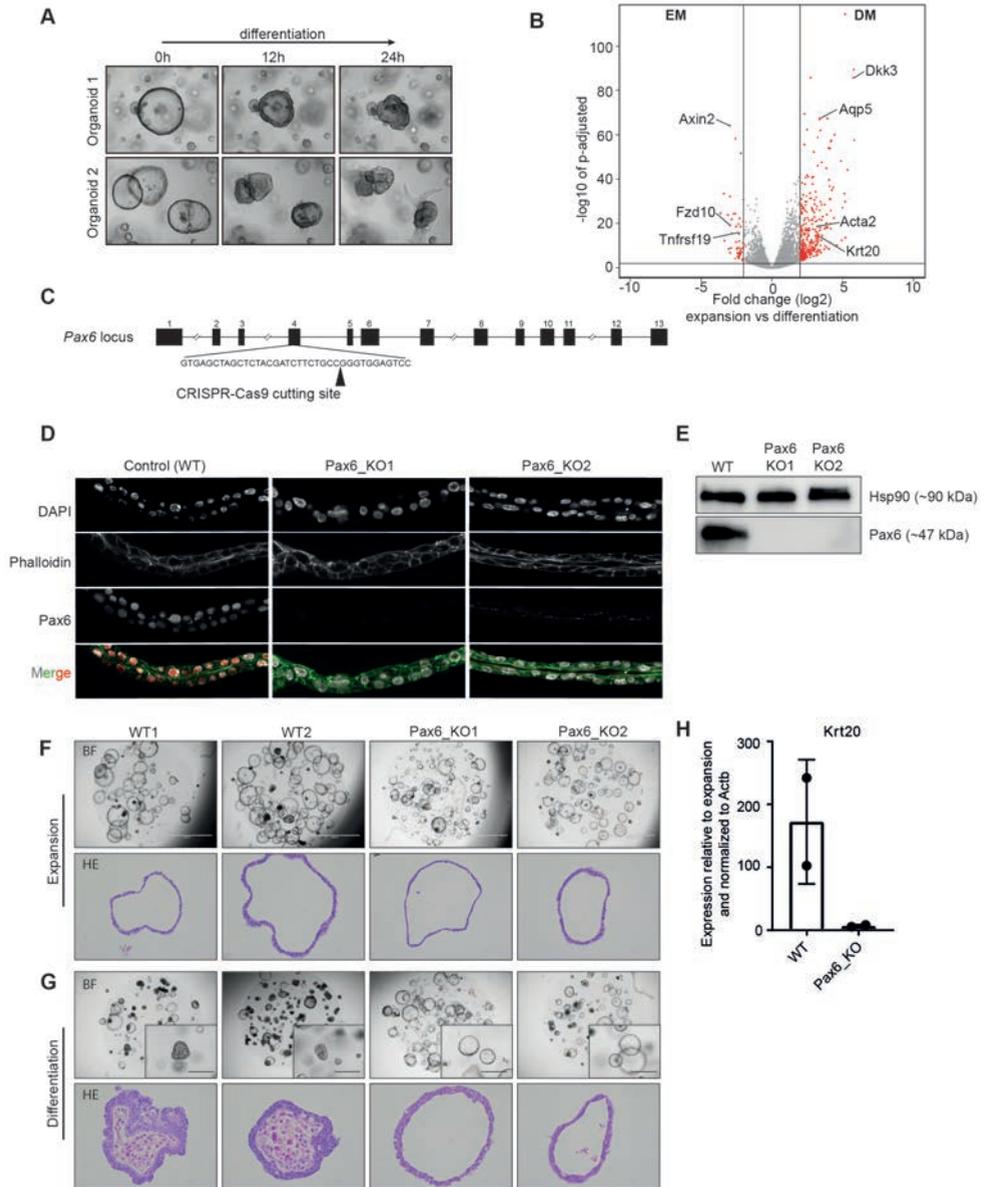
Figure 2. Differentiation of organoids depends on the expression of Pax6. (A) Representative brightfield images of organoid morphology in differentiation medium. (B) Volcano plot of differentially expressed genes in expansion and differentiation medium. (C) Schematic representation of targetting Pax6 exon 4 with CRISPR/Cas9. (D) Immunofluorescent staiting for Pax6 protein in Wiltype and Pax6-KO lines. (E) Western blot for Pax6 and Hsp90 loading control in Wildtype and Pax6-KO lines. (F-G) Brightfield images and HE staining of Wildtype and Pax6-KO lines in expansion and differentiation. (H) Expression levels of Krt20 in Wildtype and Pax6-KO lines in differentiation medium.

(**Figure 2H**). Thus, Pax6 appears to be dispensable for adult stem cell expansion but rather to be of importance in maturation of the adult lacrimal gland epithelial cells.

In recent years, single-cell sequencing technologies have revealed the cellular composition of various human organs (Regev et al. 2017). Along these lines, we performed single-cell mRNA sequencing on primary human lacrimal gland tissue. A total of 1,140 cells were sorted and processed using the SORT-seq method (Muraro et al. 2016). K-medoids-based clustering compartmentalized the 621 cells that passed the threshold (1,500 transcripts per cell) into 10 different cell clusters, as visualized by t-distributed stochastic neighbor embedding (t-SNE) (**Figure 3A**). Mapping the expression of established lacrimal gland markers that we found in murine organoids (PAX6 and AQP5) revealed two distinct resident epithelial cell populations (clusters 2, 4, 6, 7 and 10 vs cluster 8) (**Figure 3B**). Other cell types in our dataset included myoepithelium (cluster 9 expressing *ACTA2, MYLK, TPM2*), endothelial cells (cluster 1: *VWF, PECAM, CDH5*) and lymphocytes (clusters 3 and 5: *CD8A, CD2, PTPRC*) (**Figure 3C and S3A**). Previous reports describe similar cell populations in 359 sequenced cells from a mouse lacrimal gland, postnatal day 4 (Farmer et al. 2017). A more detailed analysis of the epithelial populations uncovered that clusters 2, 4, 6, 7 and 10 (n=498) consisted of ductal cells, marked by the expression of *TFCP2L1*, a transcription factor implicated in the maturation of ducts in the salivary gland and kidney (**Figure S3B**) (Yamaguchi, Yonemura, and Takada 2006). Ductal cells also expressed *TNFRSF19*, encoding TROY, a receptor we found to be enriched in murine organoids in expansion medium (**Figure 2B and S3B**). Cluster 8 (n=42) comprised acinar cells that highly expressed *PLA2G2A*, encoding a phospholipase A2 protein (**Figure S3B**). PLA2G2A is expressed in intestinal Paneth cells, serving as a host defense molecule (Ouellette 2011). Pancreatic acinar cells express group 1a of the phospholipase A2 family of proteins (Friess et al. 2001). A direct comparison (fc ≥ 1.5, p ≤ 0.05) of ductal and acinar cells revealed differential production of individual tear components. Ductal cells expressed *HTN1, C4ORF40, LTF, LYZ, PROL1, ZG16B* and *SFRP1* (**Figure 3D and 3E**). Of note, decreased levels of Secreted Frizzled Related Protein 1 (*SFRP1*) in tear fluid has been implicated in keratoconus (You et al. 2013). The gene *C4ORF40* encodes Proline-Rich Protein 27 (PRR27). While family members of this protein have been identified in saliva, a function for this particular variant in human tear fluid is unknown (Bennick 1982). Acinar cells highly expressed *LACRT, SCGB2A1, SCGB1D1, CST4, PLA2G2A, PIP, CLU, CST1, DCD* and *HMGCS2* in our dataset (**Figure 3D and 3F**). Prolactin induced protein (PIP) is a component of human milk and saliva, detected in acinar cells of the parotid gland (Mirels, Hand, and Branin 1998), believed to directly bind to bacteria (Schenkels et al. 1997).

We noticed that the expression of some tear components was unequally distributed over the 5 ductal clusters (2, 4, 6, 7 and 10). In order to assess the heterogeneity within the ductal compartment, we extracted all 498 ductal cells from the dataset and performed subclustering. This resulted in 17 ductal subclusters, expressing different tear components (**Figure S4A**). The most striking differences were observed between clusters 6 and 12. Cluster 6 comprised cells enriched for transcripts of *LACRT*, *CST4* and *SCGB1D1* sharing some of the secreted proteins found in acinar cells (**Figure S4B**). The cells in cluster 12 and 8 were positive for *WFDC2*,
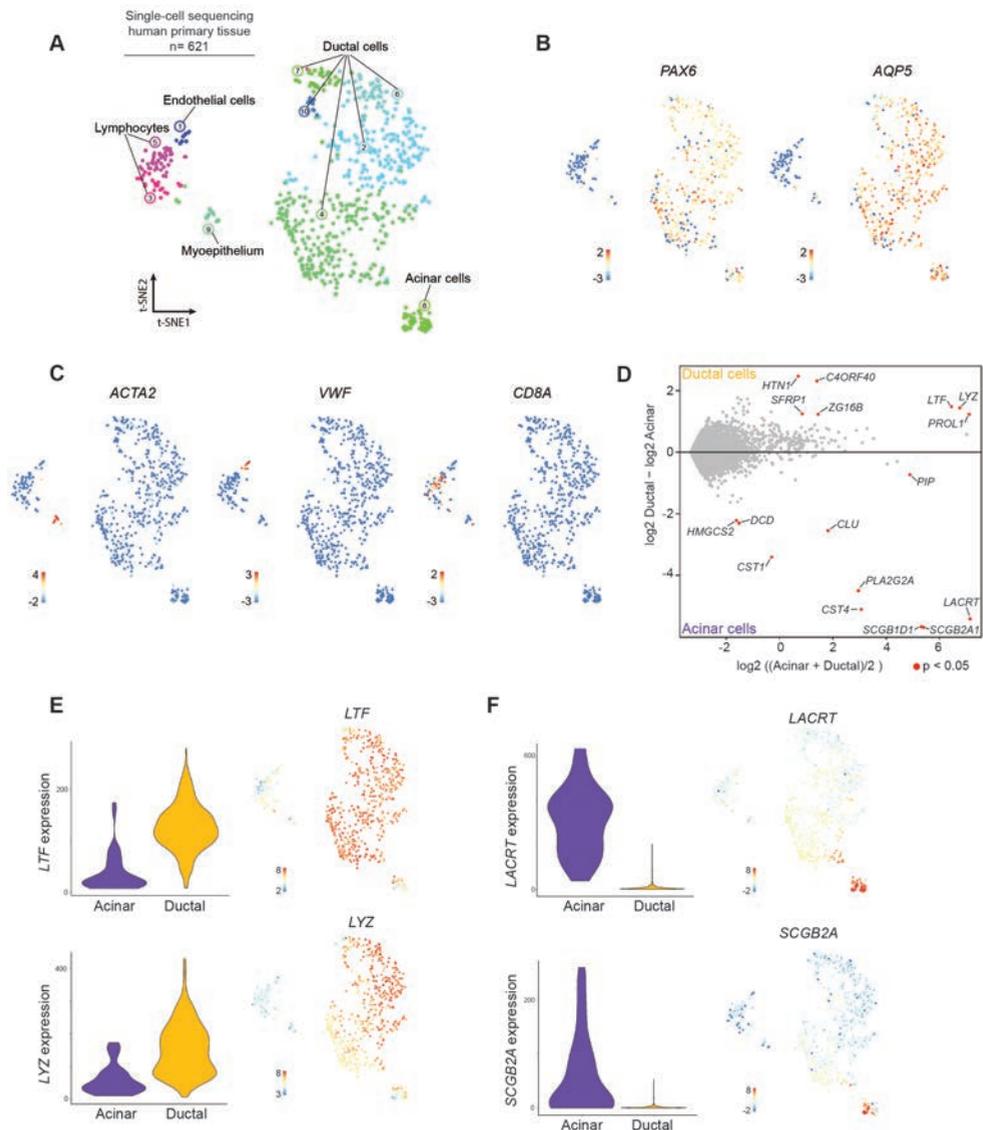
Figure 3. Single cell atlas of the human lacrimal gland reveals cellular heterogeneity in tear component expression. (A) Single cell RNA sequencing: clustering of human lacrimal gland tissue cells (n=621) visualized by t-SNE map. Colors highlight different clusters (n=10). (B-C) Expression level of selected genes in t-SNE map. Color coded logarithmic scale of transcript expression. (D) Volcano plot of differentially expressed genes comparing ductal and acinar cells from the single cell sequencing dataset. (E-F) Expression level of selected genes in t-SNE map and violin plots visualizing the expression levels in selected cells of acinar (purple) and ductal (orange) identity. Color coded logarithmic scale of transcript expression.

LCN2 and *SCGB3A1* (**Figure S4C**). Of note, these cells have much lower expression of *LTF*, *LYZ* and *AQP5* (**Figure S4D**). Secretoglobin 1D1 (*SCGB1D1*), also known as Lipophilin-A, has been identified in the proteome of human tears (Lehrer et al. 1998). Secretoglobin 3A1 (*SCGB3A1*), expressed in different ductal cells, has never been described in the context of human tears or lacrimal gland apparatus to our knowledge. Individual secretoglobin variants can be used to define secretory cell subsets in the mouse and human airways (Reynolds et al. 2002). Almost all differentially expressed genes are secreted proteins and are described to possess antimicrobial activity (McDermott 2013). Tear fluid plays an important role as an innate defense system, protecting the eye from pathogens. The loss of many of these proteins (from both ductal and acinar cells) in the tear proteome is a symptom of dry eye disease (Azharuddin et al. 2015).

We next pursued the establishment of human lacrimal gland organoids. Starting from surplus material from diagnostic lacrimal gland biopsies (n=3), we modified the mouse protocol (**Figure 4A and S5A**). Plating of dissociated tissue fragments resulted in the initial expansion of human organoids within nine days (**Figure 4B**). We observed best outgrowth and long-term expansion of cystic organoids in a growth factor medium containing R-spondin 3, Noggin, FSK, FGF10, N-acetylcysteine and A83-01 (**Figure S5B**). Human organoids could be passaged by mechanical disruption at a ratio of 1:3 every 10 days, for multiple months (>10 passages) without significant changes in growth kinetics or morphology (**Figure 4B**). Immunohistochemistry confirmed epithelial lacrimal gland identity of the organoids, matching the reference primary tissue by nuclear expression of PAX6 (**Figure 4C**).

All organoid cells were positive for Epithelial Cell Adhesion Molecule (EPCAM) (**Figure S5C**). Detection of AQP5 at the apical membrane indicates functional water movement into the lumen of the organoids (**Figure 4C**). Luminal accumulation of fluids, often following an electrochemical gradient, results in the swelling of organoids. Organoid swelling has been applied by us previously to monitor CFTR function in intestinal tract organoids of patients with cystic fibrosis (Dekkers et al. 2013). Using lacrimal gland organoid swelling as a proxy for tear fluid secretion, we tested this functional platform to monitor chemical inducers of crying. The secretion of water, electrolytes and proteins from the lacrimal gland are controlled by different neural mechanisms. Neurotransmitters (acetylcholine, VIP, Noradrenaline) released from stimulated parasympathetic and sympathetic nerves act on ductal cells, acinar cells and the myoepithelium. As expanding organoids (both mouse and human) are constitutively swollen, presumably through the presence of cAMP activator FSK in expansion medium, we first cultured organoids in a basal medium containing Advanced DMEM, B27 supplement and FGF10 (**Figure 1B and 4B**). In this basal medium, without cAMP/PKA activation, organoids shrunk resulting in an undetectable lumen. In a small screening setup, non-swollen organoids were exposed to the following compounds: cAMP/PKA activators PGE2/FSK, neurotransmitter Noradrenaline, acetylcholine receptor activator Carbachol and muscarinic receptor activator Pilocarpine. Of those, only PGE2/FSK and Noradrenaline were potent activators of secretion, as evidenced by swelling up to 300% of the initial size compared to control (**Figure 4D**). FSK directly increases cAMP levels upon administration as a cell permeable adenylyl cyclase activator. Noradrenaline requires ligand-receptor interaction. Functional adrenergic receptor
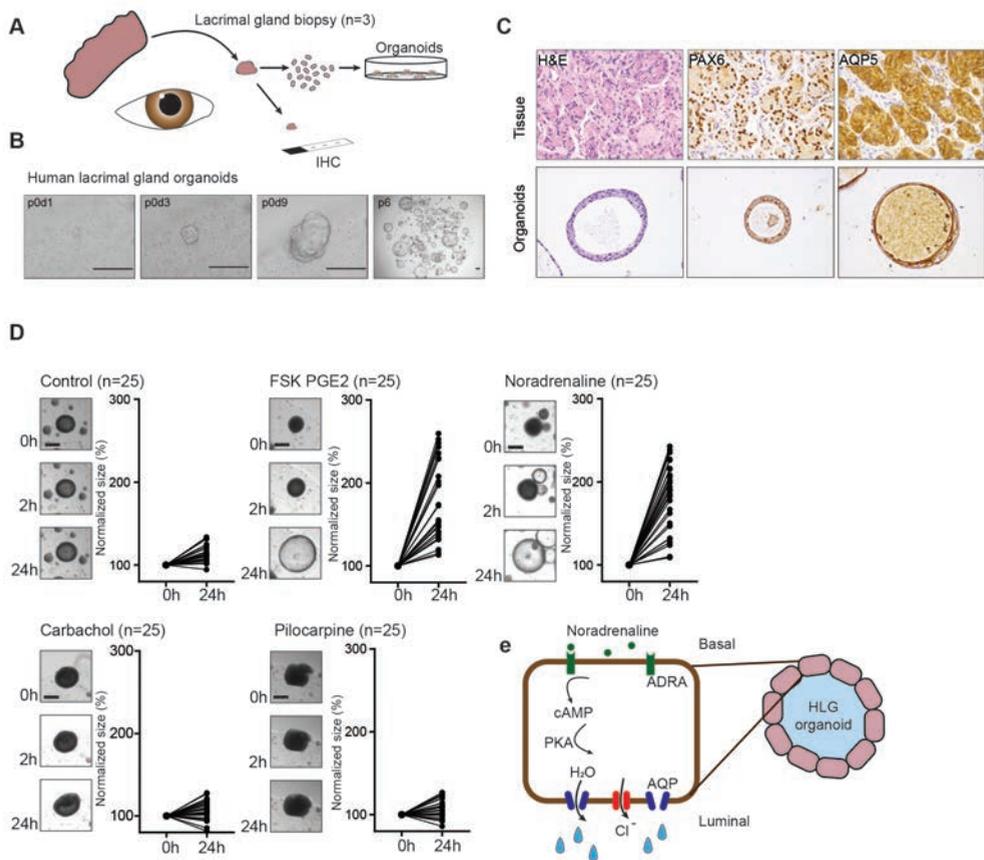
Figure 4. Human lacrimal gland organoids as a functional model for tear secretion. (A) Schematic representation of experimental setup for human lacrimal gland organoids and immunohistochemistry validation using biopsy samples. (B) Human organoid outgrowth from plating primary tissue to late-passage. (C) Immunohistochemistry of lacrimal gland markers (AQP5 and PAX6) in tissue and organoids. (D) Swelling assay of human organoids exposed to different chemical stimuli. Swelling is quantified as organoid diameter (n=25) before and 24h after exposure. Normalized to t0. (E) Schematic representation of organoid 'tearing' in response to noradrenaline.

mediated pathways alpha 1 (*Adra1a*) and beta 1 (*Adrb1*) have been described in the mouse lacrimal gland (Ding, Walcott, and Keyser 2007) (**Figure 4E**). Also, sympathetic nerves in the lacrimal gland act as a local source of noradrenaline, regulating secretion *in vivo* (Dartt 2009). Water secretion, and thereby swelling, will only take place upon functional ion channel expression, such as *CFTR, CLCN3, LRRC26* or *SLC12A4* (Ubels et al. 2006). Both carbachol and pilocarpine stimulate tear fluid secretion *in vivo* but not in organoids (Fonseca et al. 2015; Vivino et al. 1999). A number of possible explanations exist for this difference and require further investigation; (1) These agonists target receptors that are not or too lowly expressed in organoids. (2) The mechanism of action *in vivo* is indirect by stimulating the contraction of myoepithelial cells. As is described for guinea pig lacrimal gland myoepithelial cells in response

to carbachol (Satoh et al. 1997). (3) The compounds stimulate protein secretion rather than electrolyte/water secretion. This will have a limited effect on organoid swelling. The molecular mechanism for fluid secretion by lacrimal gland duct cells remains largely understudied (Katona et al. 2014).

CONCLUSION

Here, we describe an adult stem cell-based organoid protocol for the long-term expansion of mouse and human lacrimal gland. Organoids morphologically and transcriptionally recapitulate the lacrimal gland epithelium *in vivo*. CRISPR-Cas9 mediated genome editing reveals Pax6 to be crucial in the differentiation capacity of murine organoids, without affecting expansion potential or self-organization *in vitro*. Human organoids can be maintained and expanded in a very similar culture medium, activating signaling pathways also found to be essential for other organs. Human organoids are responsive to compounds that induce tearing *in vivo*. Our single-cell sequencing data demonstrate that ductal and acinar cells in the human lacrimal gland secrete a different repertoire of tear components. The organoids predominantly resemble the ductal cell state, as seen in the pancreas (Boj et al. 2015) and salivary gland (Maimets et al. 2016) organoid systems. The ductal and acinar lineages are described to have their own progenitors in the postnatal murine lacrimal gland (Farmer et al. 2017). *In vitro* transdifferentiation or optimized culture conditions for both populations may be further explored. Patient-derived organoids open up new avenues to study lacrimal gland diseases in a personalized-fashion, by following a similar biobanking approach we have used in the field of oncology (van de Wetering et al. 2015), has been done for cystic fibrosis (Berkers et al. 2019) or targeting relevant genes utilizing genome editing tools (Drost et al. 2015). Given that Sjogren's syndrome is an autoimmune disease, it would be relevant to explore co-cultures of resident immune cells with epithelial organoids, as previously illustrated with tumor-derived organoids (Dijkstra et al. 2018). In conclusion, organoids make a tractable model to better understand the lacrimal gland and study the pathologies of tearing.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

Y.P. and H.C. conceived and designed the project. Y.P., M.B-H. and M.T.B. performed experiments. M.B-H. performed analysis of RNA sequencing data. Y.E.B-H. assisted with FACS for single cell sequencing. J.v.d.V. generated salivary gland RNA sequencing data. H.B. and J.K. generated IHC data. R.K. and S.M.I. provided access to human tissue. Y.P., M.B-H. and H.C. wrote the manuscript with input from all other authors.

COMPETING FINANCIAL INTERESTS

The authors disclose no conflicts of interest.

METHODS

Patient samples

Human lacrimal gland biopsies were obtained from Rachel Kalmann and Saskia M. Imhoff at the University Medical Center Utrecht (UMCU) from patients undergoing surgery at the UMCU. The use of sample for research was approved by the medical ethical committee (TCBio) of the UMCU as protocol 18-740. Samples were kept in Advanced DMEM solution (Gibco) with Pen/Strep (Thermo-Fisher) at 4°C until further processing.

Mice

All animal experiments were performed after institutional review by the Animal Ethics Committee of the Royal Netherlands Academy of Arts and Sciences (KNAW) with project license AVD8010020151. Surplus material from female C57BL/6 was used in this study.

Lacrimal gland organoid cultures, reagents

For both mouse and human, complete or part of the gland was chopped into small pieces of approximately 1 mm using a scalpel. Muscle and connective tissue was removed and discarded as much as possible. Epithelial tissue pieces were enzymatically digested in 5 mL collagenase (Sigma-Aldrich, C9407, 1 mg/mL) with 10 µM ROCK inhibitor Y-27632 (Abmole, M1817) in AdDMEM/F12 (Gibco) for about 15 minutes shaking (120 RPM) at 37 °C. The homogenous cell suspension was pelleted and washed twice with AdDMEM/F12 prior to plating. Cells from a single murine lacrimal gland were plated in approximately 75 uL Cultrex Pathclear Reduced Growth Factor Basement Membrane Extract (BME) (3533-001, Amsbio). For human biopsies the volume of BME was determined based on the size of the final pellet. After BME solidification, culture medium was added.

Mouse expansion culture medium was based on AdDMEM/F12 (Gibco) supplemented with B27, Glutamax, HEPES, 100 U/mL Penicillin-Streptomycin (all Thermo-Fisher), 100 mg/mL Primocin (Invivogen), 1.25 mM N-acetylcysteine, 10 mM Nicotinamide (both Sigma-Aldrich) and the following growth factors: 1% Noggin conditioned medium (U-Protein Express), 1% R-spondin 3 conditioned medium (U-Protein Express), 0.5 µM A83-01 (Tocris), 1 µM PGE2 (Tocris), 1 µM FSK (Tocris) and 100 ng/mL FGF10 (Peprotech). For mouse differentiation medium we removed R-spondin, Noggin, A83-01, PGE2, Nicotinamide and N-acetylcysteine from the medium for five days.

Human expansion culture medium was based on AdDMEM/F12 (Gibco) supplemented with B27, Glutamax, HEPES, 100 U/mL Penicillin-Streptomycin (all Thermo-Fisher), 100 mg/mL Primocin (Invivogen), 1.25 mM N-acetylcysteine (Sigma-Aldrich) and the following growth

factors: 2% Noggin conditioned medium (U-Protein Express), 2% R-spondin 3 conditioned medium (U-Protein Express), 0.5 µM A83-01 (Tocris), 1 µM FSK (Tocris) and 100 ng/mL FGF10 (Peprotech).

Ten days after seeding, human and mouse organoids were removed from the BME, mechanically dissociated into small fragments using a narrowed Pasteur pipette and re-seeded in fresh BME. Passage was performed in 1:3 – 1:5 split ratio once every 7 - 10 days for at least 2 months. After initial seeding (passage 0) or after splitting the expansion medium was supplemented with 10 µM Y-27632 (Abmole). During culturing, medium was refreshed at most every three days. Images of organoid cultures were taken on EVOS FL Cell Imaging System (Thermo Fisher).

Functional swelling assay

Functional swelling of human organoids was performed with established organoid lines (> passage 5). Organoids were mechanically split as usual at a ratio of 1:3 in expansion medium. After 48 - 72 hours, the expansion medium was washed away with PBS and replaced with basal medium containing AdDMEM/F12 (Gibco) supplemented with B27, Glutamax, HEPES, 100 U/mL Penicillin-Streptomycin (all Thermo-Fisher) and 100 ng/mL FGF10 (Peprotech) to prevent pre-swelling. After 48 hours in basal medium, organoids were shrunk/solid without a lumen and ready to be exposed to the different compounds and neurotransmitters. For exposure, the basal medium was supplemented with 1 µM PGE and 1 µM FSK, 100 µM Noradrenaline (Sigma), 500 µM Carbachol (Sigma), 100 µM Pilocarpine (Sigma) or control (basal medium). Individual organoids were imaged and followed using EVOS FL Auto 2 (Thermo Fisher) before exposure and 2 hours and 24 hours after exposure. For quantification, individual organoid (n=25 per condition) diameter was measured before exposure and 24 hours after exposure using ImageJ. Swelling ratio was determined by normalizing to the diameter of the organoids before exposure.

CRISPR/Cas9 mediated Pax6 knockout

Guide RNA targeting mouse *Pax6* exon 4 (sgPax6: CCCGGCAGAAGATCGTAGAG) was cloned into the pSpCas9(BB)-2A-GFP (PX458) plasmid as described elsewhere (Ran 2013). Murine organoids were mechanically dissociated into small pieces and electroporated with 10 µg plasmid using the NEPA21 electroporator (Nepagene) as previously described (Fujii 2015). Electroporated cells were plated in BME and cultured in expansion medium. After 3 days cells were dissociated into a single cell suspension and GFP-positive cells were sorted out using the BD FACSAria (BD Biosciences). Single sorted cells that formed an organoid were picked and expanded as clonal organoid lines for further characterization. *Pax6* genotype sequence was investigated by PCR (Pax6_F1: TCCAGTGGGCAGGTTCAAAT, Pax6_R1: GGAAGGGCACTCCCGTTTAT, Pax6_F2: ATCATAGACGCGCTCCTTCC, Pax6_R2: CACAAGCGTGATGGATGCAA). If evidence was found for genomic disruption around the sgRNA targeting region the organoid was further characterized for Pax6 protein using imaging and western blot.

Western blot

Organoids were collected from the BME using cold Cell Recovery Solution (Corning) for 30 min at 4°C. Total protein was isolated using RIPA Buffer (50 mMTris-HCl pH 8.0, 150 mMNaCl, 0.1% SDS, 0.5% Na-Deoxycholate, 1% NP-40, 1X Complete protease inhibitors (Roche) and the protein extract was sonicated (how many times?)onication. Samples were prepared for gel loading by measuring the concentration with a Bradford Assay (Bio-rad reagent). Samples were loaded on a pre-cast gradient SDS-PAGE gel 12% (Biorad) and run for 1-2 h at 100 V, for subsequent transfer onto a PVDF membrane (Millipore) overnight at 100 mA and at 4°C.

Membrane was blocked with 5% milk in PBST for 1h prior to incubation with primary antibody. Antibodies were rabbit anti-Pax6 (1:1000, Biolegend 901301) and goat anti-Hsp90 (1:1000, OrigeneTA500494) as a loading control. Secondary antibodies used were swine anti-rabbit-HRP (1:1000, Dako) and rabbit anti-goat/HRP (1:1000, Dako). After washing with PBS, revealing of the membrane was performed by exposing it to a mixture of 500 μl of Luminol and 500 μl of Peroxide (Western blot detection reagent kit, GE Healthcare) and imaged using ImageQuant LAS 4000 ECL western blot imager.

Immunohistochemistry and imaging

Organoids were harvested in cell recovery solution (354253, Corning) and fixed in 4% paraformaldehyde (Sigma-Aldrich) for at least 2 hours at room temperature. Lacrimal gland tissue used for immunohistochemistry was directly embedded in 4% paraformaldehyde upon dissection and fixed for at least 2 hours at room temperature. Samples were washed and dehydrated by an increasing ethanol gradient before embedding in paraffin. Sections were cut and hydrated before staining. Hematoxylin and eosin (H&E) staining on organoids and tissue was performed as previously described (Sato et al., 2009). Antibody stainings on paraffin sections included PAX6 (Biolegend 901301), AQP5 (Origene TA307525), Ltf (Millipore 07-685) according to manufacturer's instructions. Slides were imaged using a Leica DM4000 microscope.

For immunofluorescence organoids were harvested, fixed and permeabilized using 0,2% Triton X-100. Whole-mount staining was performed overnight in 2% donkey serum using the following reagents: anti-Pax6 (Biolegend, 901301), anti-Lcn2 (R&D systems AF1757), Alexa-fluor-647 Phalloidin (1:1000, Thermo Fisher Scientific) and DAPI (Invitrogen). Secondary antibodies included Alexa Fluor 568 donkey anti-rabbit IgG (Life Technologies, A10042) and Alexa Fluor 568 donkey anti-goat IgG (Life Technologies, A11057). Organoids were imaged on a Leica SP8X microscope.

qPCR and bulk mRNA sequencing

RNA isolations of organoids and tissue for bulk RNA sequencing and qPCR were performed with RNeasy Mini Kit (Qiagen) following the manufacturer's instructions. Quantitative PCR analysis was performed using the SYBR Green and Bio-Rad systems. Changes in expression relative to beta-Actin were calculated using CFX manager software (Bio-Rad). Primers were designed using NCBI primer design tool: Krt20_F: ACAGTTCGAGAGACAGAGTCA,

Krt20_R: GGTTCTGGTAGGTGCGTCTC, Actb_F: ATGCTCCCCGGGCTGTAT, Actb_R: CATAGGAGTCCTTCTGACCCATTC.

Bulk RNA sequencing was performed with the Utrecht Sequencing Facility (USEQ). In short, polyA enriched RNA was reverse transcribed and paired-end reads mapped to the mouse genome. Expression data was analyzed using DESeq2 (Love et al., 2014).

Single cell RNA sequencing

Human lacrimal gland was dissociated with collagenase I as described above and subsequently resuspended in TrypLE Express (Gibco) pre-heated to 37°C and dissociated under repeated pipetting. When the gland was fully dissociated into single cells, samples were pelleted, washed, resuspended in FACS buffer (advanced DMEM, 10 mM Y-27632 and DAPI) and strained (35 µm). Cells were immediately sorted into 384-well plates containing ERCC spike-ins (Agilent), RT primers and dNTPs (Promega) using a BD FACSJazz (BD Biosciences). Plates were prepared using Mosquito HTS (TTPlabtech). Single-cell RNA sequencing libraries were prepared following the SORT-seq protocol (Muraro et al., 2016), which is based on the CEL-seq2 method (Hashimshony et al., 2016). In short, cells were first lysed for 5 min at 65 °C, and RT and second-strand mixes were dispensed by the Nanodrop II liquid handling platform (GC Biotech). Single-cell double-stranded cDNAs were pooled together and in vitro transcribed for linear amplification. Illumina sequencing libraries were prepared using the TruSeq small RNA primers (Illumina) and these DNA libraries were sequenced paired-end, respectively, on the Illumina NextSeq. Cells with fewer than 1,500 transcripts were filtered out. The dataset was analysed using RaceID3 (Herman et al., 2018).

ACKNOWLEDGEMENTS

COMPETING FINANCIAL INTERESTS

The authors disclose no conflicts of interest.

AUTHOR CONTRIBUTIONS

Y.P. and H.C. conceived and designed the project. Y.P., M.B-H. and M.T.B. performed experiments. M.B-H. performed analysis of RNA sequencing data. Y.E.B-H. assisted with FACS for single cell sequencing. J.v.d.V. generated salivary gland RNA sequencing data. H.B. and J.K. generated IHC data. R.K. and S.M.I. provided access to human tissue. Y.P., M.B-H. and H.C. wrote the manuscript with input from all other authors.

REFERENCES

Azharuddin, M., J. Khandelwal, H. Datta, A. K. Dasgupta, and S. O. Raja. 2015. 'Dry eye: a protein conformational disease', *Invest Ophthalmol Vis Sci*, 56: 1423-9.

Bennick, A. 1982. 'Salivary proline-rich proteins', *Mol Cell Biochem*, 45: 83-99.

Berkers, G., P. van Mourik, A. M. Vonk, E. Kruisselbrink, J. F. Dekkers, K. M. de Winter-de Groot, H. G. M. Arets, R. E. P. Marck-van der Wilt, J. S. Dijkema, M. M. Vanderschuren, R. H. J. Houwen, H. G. M. Heijerman, E. A. van de Graaf, S. G. Elias, C. J. Majoor, G. H. Koppelman, J. Roukema, M. Bakker, H. M. Janssens, R. van der Meer, R. G. J. Vries, H. C. Clevers, H. R. de Jonge, J. M. Beekman, and C. K. van der Ent. 2019. 'Rectal Organoids Enable Personalized Treatment of Cystic Fibrosis', *Cell Rep*, 26: 1701-08 e3.

Boj, S. F., C. I. Hwang, L. A. Baker, Chio, II, D. D. Engle, V. Corbo, M. Jager, M. Ponz-Sarvise, H. Tiriac, M. S. Spector, A. Gracanin, T. Oni, K. H. Yu, R. van Boxtel, M. Huch, K. D. Rivera, J. P. Wilson, M. E. Feigin, D. Ohlund, A. Handly-Santana, C. M. Ardito-Abraham, M. Ludwig, E. Elyada, B. Alagesan, G. Biffi, G. N. Yordanov, B. Delcuze, B. Creighton, K. Wright, Y. Park, F. H. Morsink, I. Q. Molenaar, I. H. Borel Rinkes, E. Cuppen, Y. Hao, Y. Jin, I. J. Nijman, C. Iacobuzio-Donahue, S. D. Leach, D. J. Pappin, M. Hammell, D. S. Klimstra, O. Basturk, R. H. Hruban, G. J. Offerhaus, R. G. Vries, H. Clevers, and D. A. Tuveson. 2015. 'Organoid models of human and mouse ductal pancreatic cancer', *Cell*, 160: 324-38.

Brito-Zeron, P., C. Baldini, H. Bootsma, S. J. Bowman, R. Jonsson, X. Mariette, K. Sivils, E. Theander, A. Tzioufas, and M. Ramos-Casals. 2016. 'Sjogren syndrome', *Nat Rev Dis Primers*, 2: 16047.

Caffery, B., E. Joyce, A. Boone, A. Slomovic, T. Simpson, L. Jones, and M. Senchyna. 2008. 'Tear lipocalin and lysozyme in Sjogren and non-Sjogren dry eye', *Optom Vis Sci*, 85: 661-7.

Chan, C. W., N. A. Wong, Y. Liu, D. Bicknell, H. Turley, L. Hollins, C. J. Miller, J. L. Wilding, and W. F. Bodmer. 2009. 'Gastrointestinal differentiation marker Cytokeratin 20 is regulated by homeobox gene CDX1', *Proc Natl Acad Sci U S A*, 106: 1936-41.

Chow, R. L., C. R. Altmann, R. A. Lang, and A. Hemmati-Brivanlou. 1999. 'Pax6 induces ectopic eyes in a vertebrate', *Development*, 126: 4213-22.

Clevers, H. 2016. 'Modeling Development and Disease with Organoids', *Cell*, 165: 1586-97.

Dartt, D. A. 2009. 'Neural regulation of lacrimal gland secretory processes: relevance in dry eye diseases', *Prog Retin Eye Res*, 28: 155-77.

de Souza, G. A., L. M. Godoy, and M. Mann. 2006. 'Identification of 491 proteins in the tear fluid proteome reveals a large number of proteases and protease inhibitors', *Genome Biol*, 7: R72.

Dekkers, J. F., C. L. Wiegerinck, H. R. de Jonge, I. Bronsveld, H. M. Janssens, K. M. de Winter-de Groot, A. M. Brandsma, N. W. de Jong, M. J. Bijvelds, B. J. Scholte, E. E. Nieuwenhuis, S. van den Brink, H. Clevers, C. K. van der Ent, S. Middendorp, and J. M. Beekman. 2013. 'A functional CFTR assay using primary cystic fibrosis intestinal organoids', *Nat Med*, 19: 939-45.

Dijkstra, K. K., C. M. Cattaneo, F. Weeber, M. Chalabi, J. van de Haar, L. F. Fanchi, M. Slagter, D. L. van der Velden, S. Kaing, S. Kelderman, N. van Rooij, M. E. van Leerdam, A. Depla, E. F. Smit, K. J. Hartemink, R. de Groot, M. C. Wolkers, N. Sachs, P. Snaebjornsson, K. Monkhorst, J. Haanen, H. Clevers, T. N. Schumacher, and E. E. Voest. 2018. 'Generation of Tumor-Reactive T Cells by Co-culture of Peripheral Blood Lymphocytes and Tumor Organoids', *Cell*, 174: 1586-98 e12.

Ding, C., B. Walcott, and K. T. Keyser. 2007. 'The alpha1- and beta1-adrenergic modulation of lacrimal gland function in the mouse', *Invest Ophthalmol Vis Sci*, 48: 1504-10.

Drost, J., W. R. Karthaus, D. Gao, E. Driehuis, C. L. Sawyers, Y. Chen, and H. Clevers. 2016. 'Organoid culture systems for prostate epithelial and cancer tissue', *Nat Protoc*, 11: 347-58.

Drost, J., R. H. van Jaarsveld, B. Ponsioen, C. Zimberlin, R. van Boxtel, A. Buijs, N. Sachs, R. M. Overmeer, G. J. Offerhaus, H. Begthel, J. Korving, M. van de Wetering, G. Schwank, M. Logtenberg, E. Cuppen, H. J. Snippert, J. P. Medema, G. J. Kops, and H. Clevers. 2015. 'Sequential cancer mutations in cultured human intestinal stem cells', *Nature*, 521: 43-7.

Farmer, D. T., S. Nathan, J. K. Finley, K. Shengyang Yu, E. Emmerson, L. E. Byrnes, J. B. Sneddon, M. T. McManus, A. D. Tward, and S. M. Knox. 2017. 'Defining epithelial cell dynamics and lineage relationships in the developing lacrimal gland', *Development*, 144: 2517-28.

Flanagan, J. L., and M. D. Willcox. 2009. 'Role of lactoferrin in the tear film', *Biochimie*, 91: 35-43.

Fonseca, B., A. Martinez-Aguila, M. Diaz-Hernandez, and J. Pintor. 2015. 'Diadenosine tetraphosphate contributes to carbachol-induced tear secretion', *Purinergic Signal*, 11: 87-93.

Friess, H., S. Shrikhande, E. Riesle, M. Kashiwagi, K. Baczako, A. Zimmermann, W. Uhl, and M. W. Buchler. 2001. 'Phospholipase A2 isoforms in acute pancreatitis', *Ann Surg*, 233: 204-12.

Fujii, M., M. Matano, K. Nanki, and T. Sato. 2015. 'Efficient genetic engineering of human intestinal organoids using electroporation', *Nat Protoc*, 10: 1474-85.

Gehring, W. J. 1996. 'The master control gene for morphogenesis and evolution of the eye', *Genes Cells*, 1: 11-5.

Halder, G., P. Callaerts, and W. J. Gehring. 1995. 'Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila', *Science*, 267: 1788-92.

Hashimshony, T., N. Senderovich, G. Avital, A. Klochendler, Y. de Leeuw, L. Anavy, D. Gennert, S. Li, K. J. Livak, O. Rozenblatt-Rosen, Y. Dor, A. Regev, and I. Yanai. 2016. 'CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq', *Genome Biol*, 17: 77.

Herman, J. S., Sagar, and D. Grun. 2018. 'FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data', *Nat Methods*, 15: 379-86.

Hirayama, M., S. B. H. Ko, T. Kawakita, T. Akiyama, S. K. Goparaju, A. Soma, Y. Nakatake, M. Sakota, N. Chikazawa-Nohtomi, S. Shimmura, K. Tsubota, and M. S. H. Ko. 2017. 'Identification of transcription factors that promote the differentiation of human pluripotent stem cells into lacrimal gland epithelium-like cells', *NPJ Aging Mech Dis*, 3: 1.

Hirayama, M., M. Ogawa, M. Oshima, Y. Sekine, K. Ishida, K. Yamashita, K. Ikeda, S. Shimmura, T. Kawakita, K. Tsubota, and T. Tsuji. 2013. 'Functional lacrimal gland regeneration by transplantation of a bioengineered organ germ', *Nat Commun*, 4: 2497.

Hodges, R. R., and D. A. Dartt. 2003. 'Regulatory pathways in lacrimal gland epithelium', *Int Rev Cytol*, 231: 129-96.

Hu, H., H. Gehart, B. Artegiani, L. Opez-Iglesias C, F. Dekkers, O. Basak, J. van Es, S. M. Chuva de Sousa Lopes, H. Begthel, J. Korving, M. van den Born, C. Zou, C. Quirk, L. Chiriboga, C. M. Rice, S. Ma, A. Rios, P. J. Peters, Y. P. de Jong, and H. Clevers. 2018. 'Long-Term Expansion of Functional Mouse and Human Hepatocytes as 3D Organoids', *Cell*, 175: 1591-606 e19.

Janssen, P. T., and O. P. van Bijsterveld. 1983. 'A simple test for lacrimal gland function: a tear lactoferrin assay by radial immunodiffusion', *Graefes Arch Clin Exp Ophthalmol*, 220: 171-4.

Jenssen, H., and R. E. Hancock. 2009. 'Antimicrobial properties of lactoferrin', *Biochimie*, 91: 19-29.

Katona, M., E. Vizvari, L. Nemeth, A. Facsko, V. Venglovecz, Z. Rakonczay, Jr., P. Hegyi, and E. Toth-Molnar. 2014. 'Experimental evidence of fluid secretion of rabbit lacrimal gland duct epithelium', *Invest Ophthalmol Vis Sci*, 55: 4360-7.

Kimoto, H., K. Sato, F. Nodari, S. Haga, T. E. Holy, and K. Touhara. 2007. 'Sex- and strain-specific expression and vomeronasal activity of mouse ESP family peptides', *Curr Biol*, 17: 1879-84.

Lehrer, R. I., G. Xu, A. Abduragimov, N. N. Dinh, X. D. Qu, D. Martin, and B. J. Glasgow. 1998. 'Lipophilin, a novel heterodimeric protein of human tears', *FEBS Lett*, 432: 163-7.

Lin, H., Y. Liu, and S. Yiu. 2019. 'Three Dimensional Culture of Potential Epithelial Progenitor Cells in Human Lacrimal Gland', *Transl Vis Sci Technol*, 8: 32.

Love, M. I., W. Huber, and S. Anders. 2014. 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol*, 15: 550.

Maimets, M., C. Rocchi, R. Bron, S. Pringle, J. Kuipers, B. N. Giepmans, R. G. Vries, H. Clevers, G. de Haan, R. van Os, and R. P. Coppes. 2016. 'Long-Term In Vitro Expansion of Salivary Gland Stem Cells Driven by Wnt Signals', *Stem Cell Reports*, 6: 150-62.

Makarenkova, H. P., M. Ito, V. Govindarajan, S. C. Faber, L. Sun, G. McMahon, P. A. Overbeek, and R. A. Lang. 2000. 'FGF10 is an inducer and Pax6 a competence factor for lacrimal gland development', *Development*, 127: 2563-72.

McDermott, A. M. 2013. 'Antimicrobial compounds in tears', *Exp Eye Res*, 117: 53-61.

Mirels, L., A. R. Hand, and H. J. Branin. 1998. 'Expression of gross cystic disease fluid protein-15/Prolactin-inducible protein in rat salivary glands', *J Histochem Cytochem*, 46: 1061-71.

Muraro, M. J., G. Dharmadhikari, D. Grun, N. Groen, T. Dielen, E. Jansen, L. van Gurp, M. A. Engelse, F. Carlotti, E. J. de Koning, and A. van Oudenaarden. 2016. 'A Single-Cell Transcriptome Atlas of the Human Pancreas', *Cell Syst*, 3: 385-94 e3.

Murube, J. 2009. 'Basal, reflex, and psycho-emotional tears', *Ocul Surf*, 7: 60-6.

Ouellette, A. J. 2011. 'Paneth cell alpha-defensins in enteric innate immunity', *Cell Mol Life Sci*, 68: 2215-29.

Perumal, N., S. Funke, N. Pfeiffer, and F. H. Grus. 2016. 'Proteomics analysis of human tears from aqueous-deficient and evaporative dry eye patients', *Sci Rep*, 6: 29629.

Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang. 2013. 'Genome engineering using the CRISPR-Cas9 system', *Nat Protoc*, 8: 2281-308.

Regev, A., S. A. Teichmann, E. S. Lander, I. Amit, C. Benoist, E. Birney, B. Bodenmiller, P. Campbell, P. Carninci, M. Clatworthy, H. Clevers, B. Deplancke, I. Dunham, J. Eberwine, R. Eils, W. Enard, A. Farmer, L. Fugger, B. Gottgens, N. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg, J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea, G. Nolan, D. Pe'er, A. Phillipakis, C. P. Ponting, S. Quake, W. Reik, O. Rozenblatt-Rosen,

J. Sanes, R. Satija, T. N. Schumacher, A. Shalek, E. Shapiro, P. Sharma, J. W. Shin, O. Stegle, M. Stratton, M. J. T. Stubbington, F. J. Theis, M. Uhlen, A. van Oudenaarden, A. Wagner, F. Watt, J. Weissman, B. Wold, R. Xavier, N. Yosef, and Participants Human Cell Atlas Meeting. 2017. 'The Human Cell Atlas', *Elife*, 6.

Reynolds, S. D., P. R. Reynolds, G. S. Pryhuber, J. D. Finder, and B. R. Stripp. 2002. 'Secretoglobins SCGB3A1 and SCGB3A2 define secretory cell subsets in mouse and human airways', *Am J Respir Crit Care Med*, 166: 1498-509.

Sato, T., D. E. Stange, M. Ferrante, R. G. Vries, J. H. Van Es, S. Van den Brink, W. J. Van Houdt, A. Pronk, J. Van Gorp, P. D. Siersema, and H. Clevers. 2011. 'Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium', *Gastroenterology*, 141: 1762-72.

Sato, T., R. G. Vries, H. J. Snippert, M. van de Wetering, N. Barker, D. E. Stange, J. H. van Es, A. Abo, P. Kujala, P. J. Peters, and H. Clevers. 2009. 'Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche', *Nature*, 459: 262-5.

Satoh, Y., K. Sano, Y. Habara, and T. Kanno. 1997. 'Effects of carbachol and catecholamines on ultrastructure and intracellular calcium-ion dynamics of acinar and myoepithelial cells of lacrimal glands', *Cell Tissue Res*, 289: 473-85.

Schenkels, L. C., E. Walgreen-Weterings, L. C. Oomen, J. G. Bolscher, E. C. Veerman, and A. V. Nieuw Amerongen. 1997. 'In vivo binding of the salivary glycoprotein EP-GP (identical to GCDFP-15) to oral and non-oral bacteria detection and identification of EP-GP binding species', *Biol Chem*, 378: 83-8.

Schutgens, F., M. B. Rookmaaker, T. Margaritis, A. Rios, C. Ammerlaan, J. Jansen, L. Gijzen, M. Vormann, A. Vonk, M. Viveen, F. Y. Yengej, S. Derakhshan, K. M. de Winter-de Groot, B. Artegiani, R. van Boxtel, E. Cuppen, A. P. A. Hendrickx, M. M. van den Heuvel-Eibrink, E. Heitzer, H. Lanz, J. Beekman, J. L. Murk, R. Masereeuw, F. Holstege, J. Drost, M. C. Verhaar, and H. Clevers. 2019. 'Tubuloids derived from human adult kidney and urine for personalized disease modeling', *Nat Biotechnol*, 37: 303-13.

Smith, R. S., A. Zabaleta, T. Kume, O. V. Savinova, S. H. Kidson, J. E. Martin, D. Y. Nishimura, W. L. Alward, B. L. Hogan, and S. W. John. 2000. 'Haploinsufficiency of the transcription factors FOXC1 and FOXC2 results in aberrant ocular development', *Hum Mol Genet*, 9: 1021-32.

Stange, D. E., B. K. Koo, M. Huch, G. Sibbel, O. Basak, A. Lyubimova, P. Kujala, S. Bartfeld, J. Koster, J. H. Geahlen, P. J. Peters, J. H. van Es, M. van de Wetering, J. C. Mills, and H. Clevers. 2013. 'Differentiated Troy+ chief cells act as reserve stem cells to generate all lineages of the stomach epithelium', *Cell*, 155: 357-68.

Tiwari, S., M. J. Ali, M. M. Balla, M. N. Naik, S. G. Honavar, V. A. Reddy, and G. K. Vemuganti. 2012. 'Establishing human lacrimal gland cultures with secretory function', *PLoS One*, 7: e29458.

Tsubota, K., S. Hirai, L. S. King, P. Agre, and N. Ishida. 2001. 'Defective cellular trafficking of lacrimal gland aquaporin-5 in Sjogren's syndrome', *Lancet*, 357: 688-9.

Ubels, J. L., H. M. Hoffman, S. Srikanth, J. H. Resau, and C. P. Webb. 2006. 'Gene expression in rat lacrimal gland duct cells collected using laser capture microdissection: evidence for K+ secretion by duct cells', *Invest Ophthalmol Vis Sci*, 47: 1876-85.

van de Wetering, M., H. E. Francies, J. M. Francis, G. Bounova, F. Iorio, A. Pronk, W. van Houdt, J. van Gorp, A. Taylor-Weiner, L. Kester, A. McLaren-Douglas, J. Blokker, S. Jaksani, S. Bartfeld, R. Volckman, P. van Sluis, V. S. Li, S. Seepo, C. Sekhar Pedamallu, K. Cibulskis, S. L. Carter, A. McKenna, M. S. Lawrence, L. Lichtenstein, C. Stewart, J. Koster, R. Versteeg, A. van Oudenaarden, J. Saez-Rodriguez, R. G. Vries,

G. Getz, L. Wessels, M. R. Stratton, U. McDermott, M. Meyerson, M. J. Garnett, and H. Clevers. 2015. 'Prospective derivation of a living organoid biobank of colorectal cancer patients', *Cell*, 161: 933-45.

Vivino, F. B., I. Al-Hashimi, Z. Khan, F. G. LeVeque, P. L. Salisbury, 3rd, T. K. Tran-Johnson, C. C. Muscoplat, M. Trivedi, B. Goldlust, and S. C. Gallagher. 1999. 'Pilocarpine tablets for the treatment of dry mouth and dry eye symptoms in patients with Sjogren syndrome: a

randomized, placebo-controlled, fixed-dose, multicenter trial. P92-01 Study Group', *Arch Intern Med*, 159: 174-81.

Yamaguchi, Y., S. Yonemura, and S. Takada. 2006. 'Grainyhead-related transcription factor is required for duct maturation in the salivary gland and the kidney of the mouse', *Development*, 133: 4737-48.

You, J., C. Hodge, L. Wen, J. W. McAvoy, M. C. Madigan, and G. Sutton. 2013. 'Tear levels of SFRP1 are significantly reduced in keratoconus patients', *Mol Vis*, 19: 509-xxx.
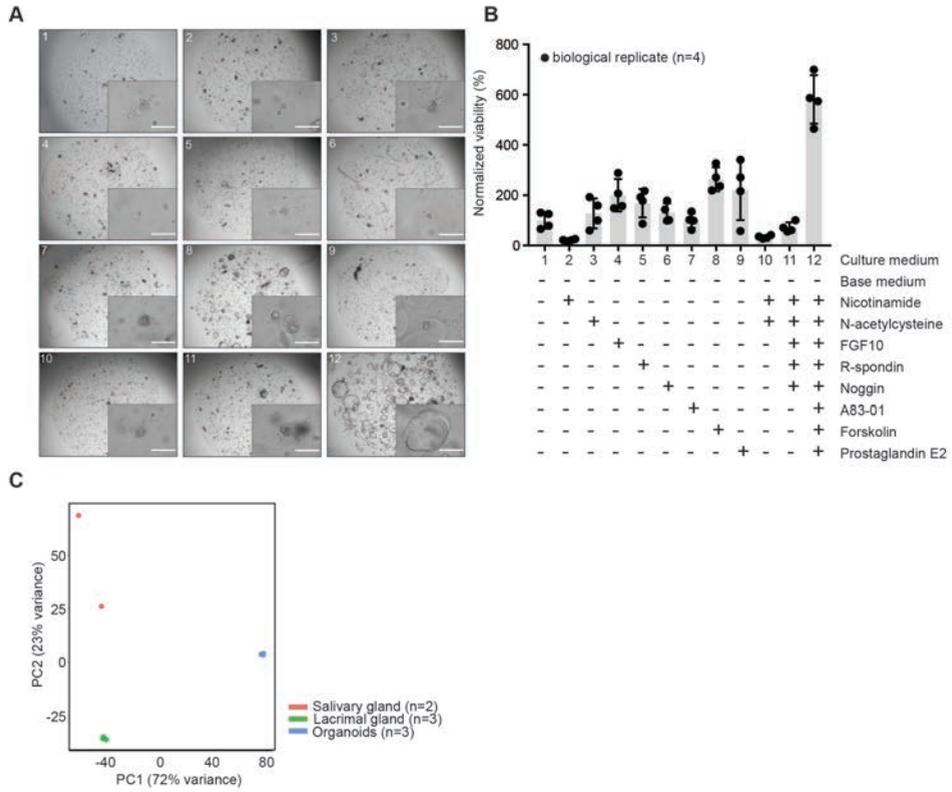
Supplementary Figure 1. Culture conditions for murine lacrimal gland organoids. (A) Representative images of organoid outgrowth from primary mouse lacrimal gland tissue in 12 different conditions (mentioned in b). (B) Quantification of (a) using Cell Titer Glo. Scheme of 12 media compositions. (C) PCA plot of bulk RNA sequencing lacrimal gland tissue, organoids and salivary gland tissue.
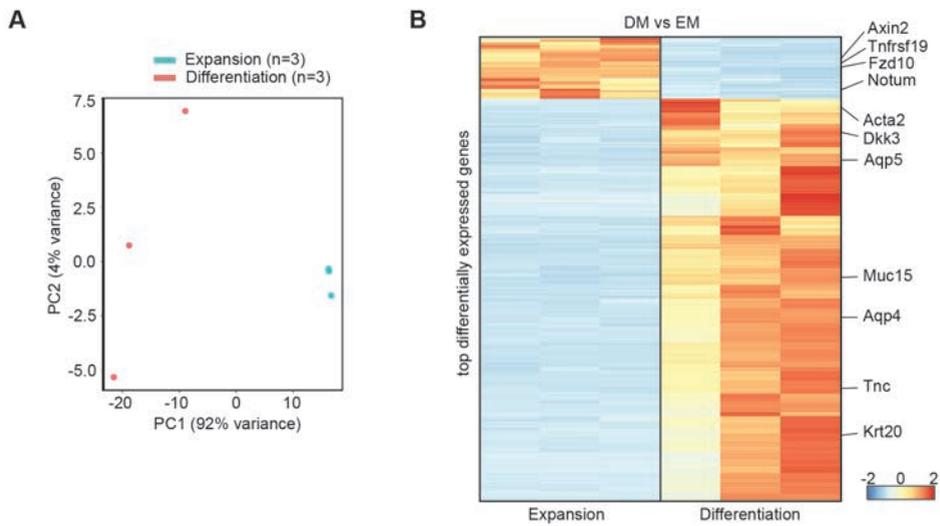
Supplementary Figure 2. Murine lacrimal gland organoids change gene expression signature upon growth factor withdrawal. (A) PCA plot of bulk mRNA sequencing of organoids in differentiation and expansion medium. (B) Heatmap of differentially expressed genes in differentiation and expansion medium.
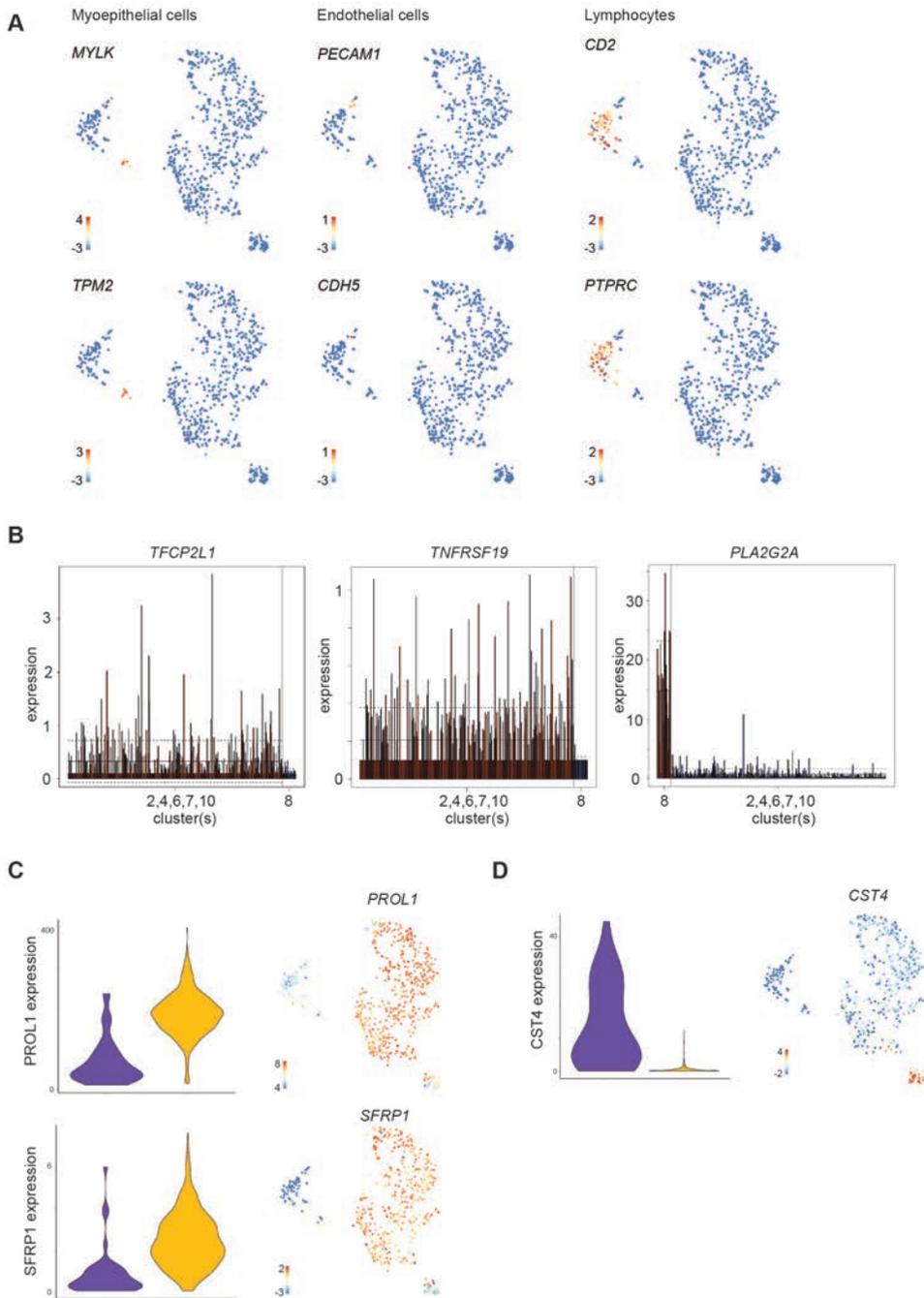
Supplementary Figure 3. Single cell transcriptome analysis of the human lacrimal gland. (A) Expression level of selected genes in t-SNE map. (B) Bar plots of expression levels (y-axis) of selected genes in epithelial cell (x-axis) clusters. (C) Expression level of selected genes in t-SNE map and violin plots visualizing the expression levels in selected cells of acinar (purple) and ductal (orange) identity. Color coded logarithmic scale of transcript expression.
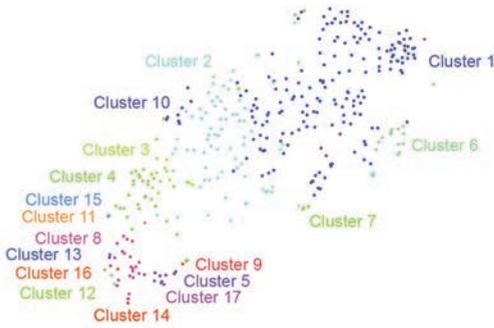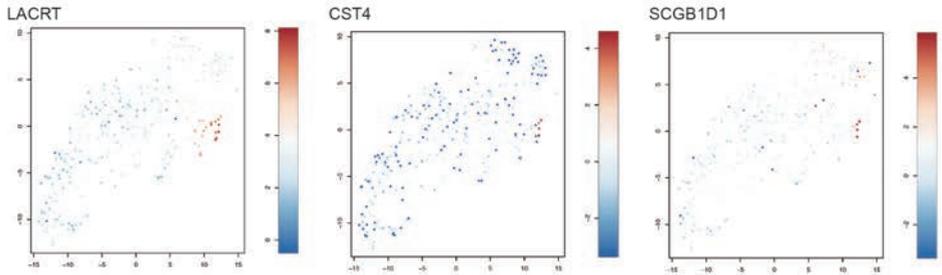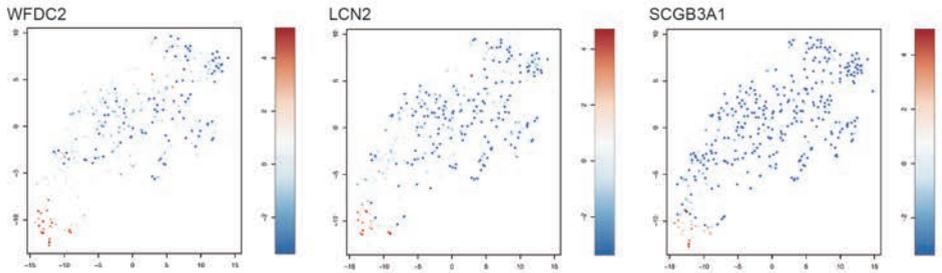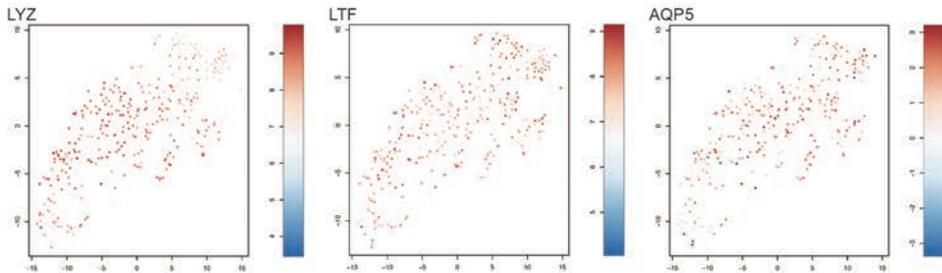
Supplementary Figure 4. Subclustering reveals ductal cell heterogeneity. (A) Subclustering of all ductal cells into 17 different clusters, as visualized by t-SNE map.(B-D) Expression level of selected genes differentially expressed in ductal subclusters, visualized in t-SNE map (color coded logarithmic scale of transcript expression).
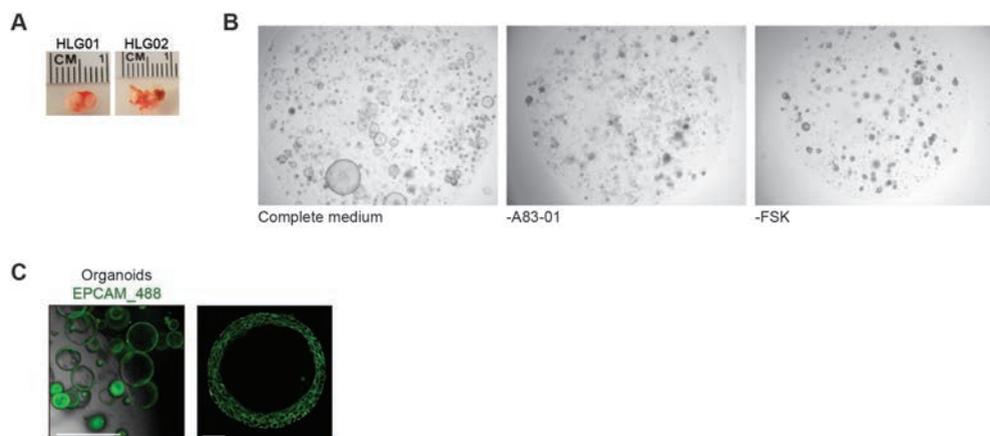
Supplementary Figure 5. Culture conditions for human lacrimal gland organoids. (A) Representative image of lacrimal gland biopsy material used for organoid culture. Before tissue dissociation and plating. (B) Brightfield images of organoid outgrowth in complete medium or medium without A83-01 or FSK. (C) Human organoids stained for EPCAM (in green).

SNAKE VENOM GLAND ORGANOIDS

4
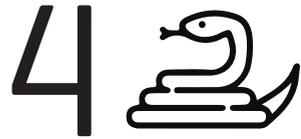
Yorick Post[1,2,13], Jens Puschhof[1,2,13], Joep Beumer[1,2,13], Harald M. Kerkkamp[3],
Merijn A.G. de Bakker[3], Julien Slagboom[4], Buys de Barbanson[1,2], Nienke R. Wevers[5],
Xandor M. Spijkers[5,6], Thomas Olivier[5], Taline D. Kazandjian[7], Stuart Ainsworth[7],
Carmen Lopez Iglesias[8], Willine J. van de Wetering[1,8], Maria C. Heinz[2,9], Ravian L van Ineveld[2,10],
Regina G.D.M. van Kleef[11], Harry Begthel[1,2], Jeroen Korving[1,2], Yotam E. Bar-Ephraim[1,2],
Walter Getreuer[12], Anne C. Rios[2,10], Remco H. S. Westerink[11], Hugo J. G. Snippert[2,9],
Alexander van Oudenaarden[1,2], Peter J. Peters[8], Freek J. Vonk[3], Jeroen Kool[4],
Michael K. Richardson[3], Nicholas R. Casewell[7] and Hans Clevers[1,2,10,14,*]

[1] Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences (KNAW) and UMC Utrecht, 3584 CT Utrecht, The Netherlands.
[2] Oncode Institute, Hubrecht Institute, 3584 CT Utrecht, The Netherlands.
[3] Institute of Biology Leiden, Department of Animal Science and Health, 2333 BE Leiden, The Netherlands.
[4] Division of BioAnalytical Chemistry, Department of Chemistry and Pharmaceutical Sciences, Vrije Universiteit Amsterdam, 1081 LA Amsterdam, The Netherlands.
[5] Mimetas BV, Organ-on-a-chip Company, 2333 CH Leiden, The Netherlands.
[6] Department of Translational Neuroscience, Utrecht University Medical Center, 3584 CG, Utrecht, The Netherlands.
[7] Centre for Snakebite Research & Interventions, Parasitology Department, Liverpool School of Tropical Medicine, Liverpool, L3 5QA, UK.
[8] The Maastricht Multimodal Molecular Imaging institute, Maastricht University, 6229 ER Maastricht, The Netherlands.
[9] Molecular Cancer Research, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht University, 3584 CX Utrecht, The Netherlands.
[10] The Princess Maxima Center for Pediatric Oncology, 3584 CS Utrecht, The Netherlands.
[11] Neurotoxicology Research Group, Division of Toxicology, Institute for Risk Assessment Sciences (IRAS), Utrecht University, 3584 CL Utrecht, The Netherlands.
[12] Serpo, 2288 ED Rijswijk, The Netherlands.
[13] Co-first author
[14] Lead Contact
* Correspondence: h.clevers@hubrecht.eu

ABSTRACT

Wnt dependency and Lgr5 expression define multiple mammalian epithelial stem cell types. Under defined growth factor conditions, such Adult Stem Cells (ASCs) grow as 3D organoids that recapitulate essential features of the pertinent epithelium. Here, we establish long-term expanding venom gland organoids from several snake species. The newly assembled transcriptome of the Cape coral snake reveals that organoids express high levels of toxin transcripts. Single cell RNA sequencing of both organoids and primary tissue identifies distinct venom-expressing cell types as well as proliferative cells expressing homologs of known mammalian stem cell markers. A hard-wired regional heterogeneity in the expression of individual venom components is maintained in organoid cultures. Harvested venom peptides reflect crude venom composition and display biological activity. This study extends organoid technology to reptilian tissues and describes an experimentally tractable model system representing the snake venom gland.

Keywords

INTRODUCTION

Snakebite envenoming is a neglected tropical disease estimated to be responsible for >100,000 deaths worldwide each year (Gutierrez et al., 2017). Upon injection through specialized fangs, venomous snakes primarily use their venom to immobilize prey. For numerous snake species, the venom delivered by a single bite is enough to kill prey that is many times larger than the snake. During vertebrate evolution, the snake venom gland has been adapted from the salivary gland (Kochva 1987). Secretory columnar cells in the gland epithelium secrete a complex mixture of peptides and proteins, stored in the lumen and channeled to the fangs through connecting ducts. Venom toxins are classified into distinct protein families with diverse modes-of-action (Fry et al., 2009), which makes them a rich source for drugs targeting human proteins (Clark et al., 2019). Snake venoms differ greatly in toxin composition and the architecture of the machinery involved in its production and release. Venom from viperid snakes (such as the rattlesnake) mainly consists of hemotoxic enzymes (Ainsworth et al., 2018), whereas elapids (such as the king cobra) typically produce smaller peptides with neurotoxic effects (Vonk et al., 2013). Three-finger toxins (3FTx) constitute a diverse and highly expressed class of toxins that mainly exert a neurotoxic effect through interaction with acetylcholine receptors (Tsetlin 2015). Kunitz-type inhibitors (KUN) also exert neurotoxic activities, though some are anticoagulant (Harvey 2001; Millers et al., 2009). Phospholipase A2 (PLA2) proteins can act as myotoxins or neurotoxins and display antiplatelet activities (Guitiérrez and Lomonte 2013), whereas cysteine-rich secretory proteins (CRISP) block smooth muscle contraction (Yamazaki and Morita 2004). Finally, snake venom metalloproteinases (SVMP), L-amino acid oxidases (LAAO) and C-type lectins (CTL) mostly disrupt blood coagulation (Slagboom et al., 2017; Izidoro et al., 2014). Much still needs to be learned about toxin production- and release-cycles, heterogeneity of venom-producing cells and factors influencing venom composition.

Organoids are defined as self-organizing 3D structures that can be grown from stem cells and that recapitulate essential features of the tissue under study (Clevers 2016). We originally showed that a serum-free medium containing R-spondin, EGF and Noggin suffices to support the growth of mouse Lgr5+ intestinal ASCs into ever-expanding epithelial organoids. These 'mini-guts' contain all known cell types of the gut lining (Sato et al., 2009), observations that were extended through single cell RNA sequencing (Grün et al., 2015; Beumer et al., 2018, Gehart et al., 2019). Subsequently, similar R-spondin-based protocols have been reported for a wide diversity of healthy and diseased mammalian epithelia (Artegiani and Clevers 2018). Of interest to the current study, Coppes and colleagues have demonstrated the feasibility of growing mammalian salivary gland organoids (Maimets et al., 2016).

Little is known about the biology of adult stem cells in reptiles. Different short-term culture systems have been described for snake venom glands (Sells et al., 1989; Carneiro et al., 2006), yet long-term cultures capturing structural and functional properties of the snake venom gland have not been developed. Here, we aim to establish long-term culture conditions for functional snake venom gland epithelium using R-spondin-based organoid technology.

RESULTS

Epithelial organoid cultures derived from snake venom glands

Venom glands from nine different snake species, representing members of both the Elapidae (*Naja pallida, Naja annulifera, Naja nivea, Naja atra and Aspidelaps lubricus cowlesi*) and Viperidae (*Echis ocellatus, Deinagkistrodon acutus, Crotalus atrox and Bitis arietans*) families, were dissociated and embedded in basement membrane extract (BME) at 32°C (**Figure 1A**). Supplying a medium containing a 'generic' mammalian organoid cocktail (see below) resulted in initial expansion of organoids for all species (**Figure 1B and S1A**) (some were subsequently lost due to bacterial contamination, insensitive to the antibiotics used). Passaging yielded expanding organoids that histologically resembled the original gland epithelium (**Figure 1C**). The venom gland organoids could be expanded optimally using R-spondin (the Wnt signal-amplifying ligand of Lgr5 (de Lau et al., 2011)), the BMP inhibitor Noggin, EGF, the small molecule TGF beta inhibitor A83-01, PGE2 and FGF10 (**Figure 1D and S1B**). This 'expansion medium' controls the same cellular signaling pathways that are required for mammalian epithelial organoids (Fatehullah et al., 2016). Reptilian organoid expansion was only successful at lower temperatures (**Figure S1C**), consistent with the average body temperature of poikilotherm species such as snakes. When cultured at 37°C, a heat shock response (*HSPA8*) was observed within two hours, that is described to preclude cell division and consequently prohibited organoid growth (Richter et al., 2010) (**Figure S1D**). The proliferating cells self-organized into cystic spheres lined by a simple polarized epithelium (**Figure 1E and Video S1**), while budding of organoids was occasionally observed. Under these conditions, organoids expanded exponentially for > 20 passages without significant changes in growth kinetics or morphology (**Figure 1B**). As demonstrated previously for mammalian organoids (Bolhaqueiro et al., 2018), a lentiviral construct encoding histone 2B coupled to a fluorescent protein (H2B-RFP) allowed visualization of the organoid chromatin (**Figure S1E**).

Newly assembled transcriptome reveals high toxin expression in organoids

Simultaneous withdrawal of all growth factors (with the exception of PGE2 to maintain cystic organoids) for 7 days resulted in less-proliferative organoids (**Figure 2A**). These organoids contained highly polarized cells with secretory vesicles, described previously as the main producers of venom in the snake venom gland (Mackessy 1991) (**Figure 2B**). Furthermore, ciliated cells could occasionally be observed by electron-microscopy (**Figure S1F**). High-speed imaging captured functionally beating cilia in a subset of organoids (**Video S2**). Based on these features of apparently mature and functional cell types, we defined this cultured condition as our 'differentiation' cocktail.

Snake venom contains dozens to hundreds of different bioactive compounds (Fox and Serrano 2008). No annotated genome exists for *A. l. cowlesi*. To gain deeper insight into the expression of individual toxin-encoding genes, we assembled a *de novo* transcriptome of *A. l. cowlesi* using Trinity (Haas et al., 2013) (**Figure 2C**). Libraries of contigs (311,948) generated
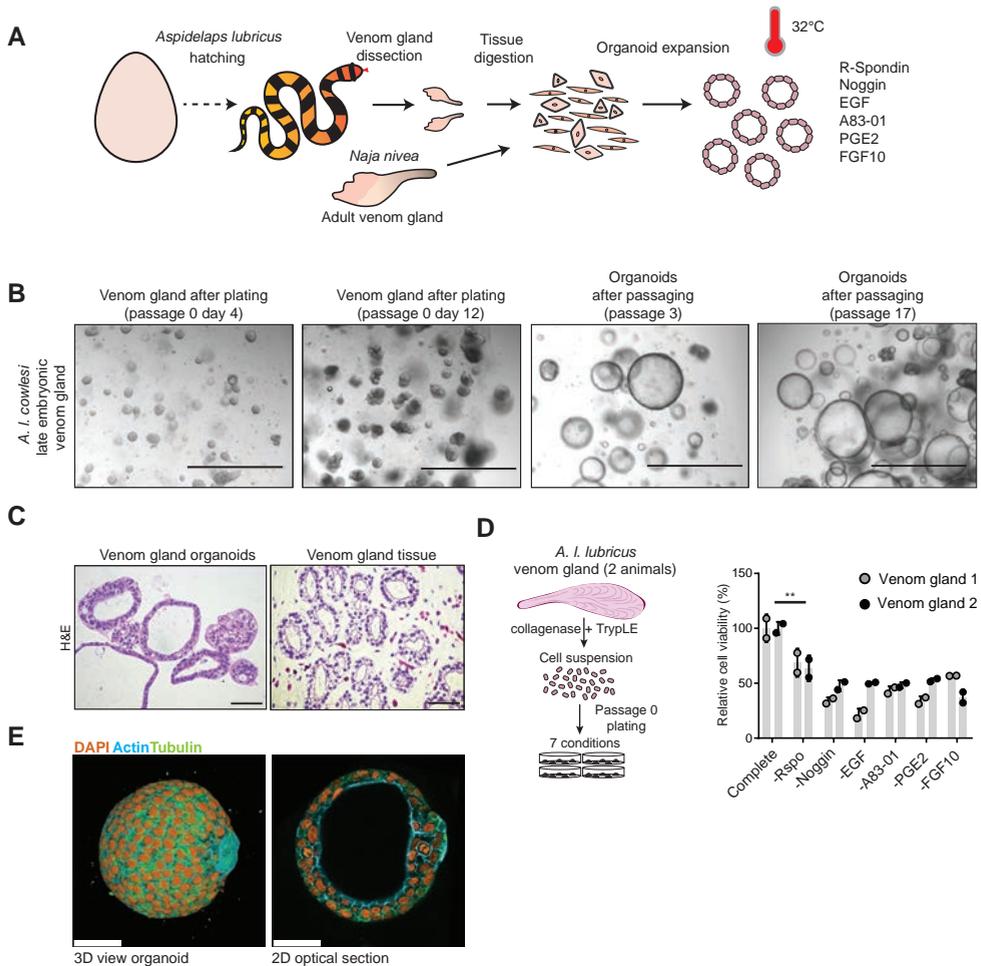
Figure 1. Establishment of organoid culture conditions for snake venom gland, see also Figure S1. (A) Schematic representation of the derivation of venom gland organoids from late embryonic (~7–2 days before hatching) *A. l. cowlesi* (n=7) (See also Figure S1A). (**B**) Time course of organoid expansion after seeding of cells from a single *A. l. cowlesi* venom gland in BME (passage 0) until passage 17. Scale bars, 1000 μm. (**C**) Haematoxylin and eosin (H&E) stain of late embryonic *A. l. cowlesi* venom gland and organoids. Scale bars, 50 μm. (**D**) Schematic representation of medium component dropout screen on primary tissue outgrowth. Quantification of relative cell viability per condition after 14 days, normalized to complete expansion medium (See also Figure S1B). Data points represent biological replicates. (**E**) Immunofluorescent staining of organoid for DNA (DAPI), Tubulin (green) and Actin (blue). Scale bars, 50 μm

from late-embryonic liver, pancreas and venom gland as well as from venom gland organoids were used to determine gene-expression levels for each of the three tissues as well as the venom gland organoids. The organoid transcriptome showed toxins to represent the dominant class of expressed genes, while homologs of established markers of mammalian liver and pancreas were restricted to their corresponding organ (**Figure 2D**). Toxin expression in venom gland
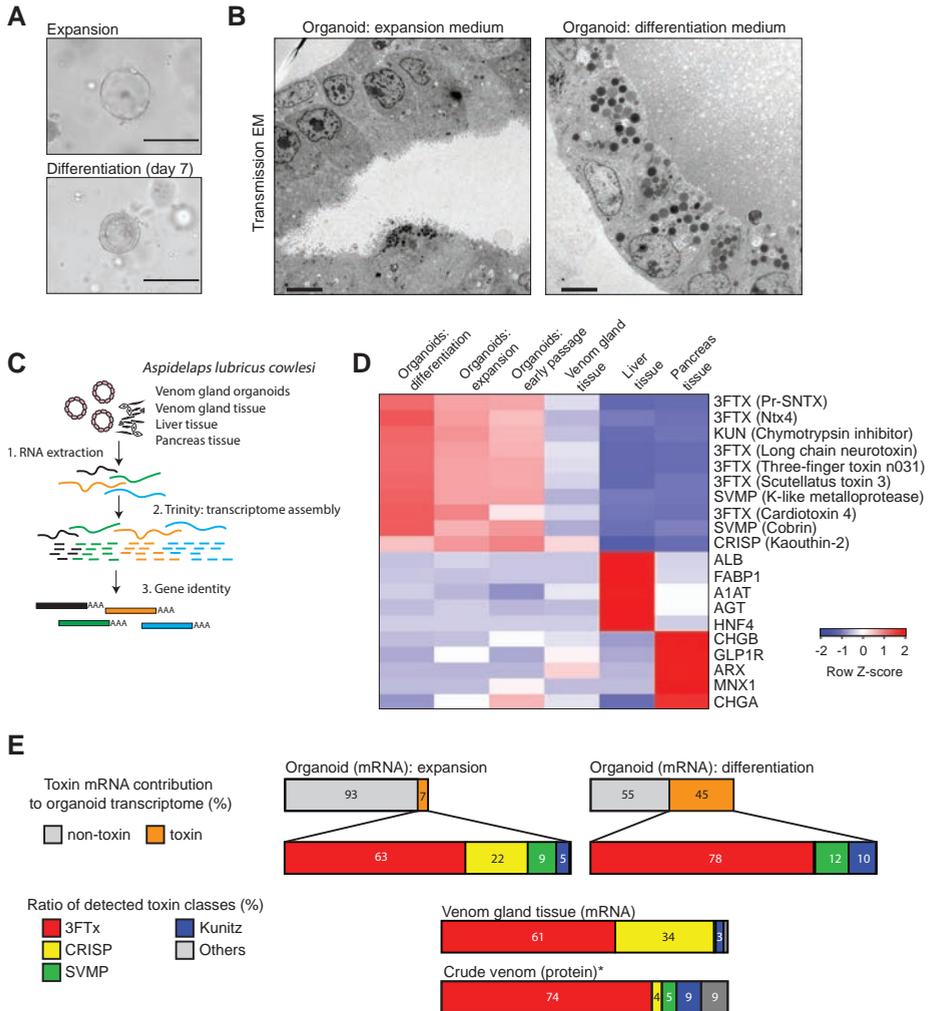
Figure 2. Venom gland organoids express a near-normal spectrum of toxins, see also Figure S1. (A) Representative bright-field images of organoids after seven days expansion medium or differentiation medium. Scale bars, 200 μm. **(B)** Transmission electron microscopy (TEM) of organoids in expansion and differentiation medium shows simple epithelial cells and polarized exocrine cells with secretory vesicles oriented towards the lumen. Scale bars, 5 μm. **(C)** Schematic representation of *de novo* transcriptome assembly using Trinity. Input contained late embryonic *A. l. cowlesi* RNA isolated from venom gland tissue, pancreas tissue, liver tissue and three venom gland organoid samples (early passage expansion, late passage expansion, late passage differentiation). **(D)** Heatmap of organoid and tissue gene expression determined by mRNA sequencing mapped on *de novo* transcriptome. Highlighted are highly expressed toxin classes (*3FTX, KUN, SVMP, CRISP*), as well as liver and pancreas markers. **(E)** Contribution of toxin-encoding genes (orange) and non-toxin genes (grey) to the transcriptome per sample, and the contribution of toxin classes in expansion medium and differentiation medium compared to venom gland tissue mRNA and venom proteome (*dataset from Whiteley and Casewell et al., 2019) (3FTx = three-finger toxin, CRISP = cysteine-rich secretory protein, SVMP = snake venom metalloproteinase, Kunitz = Kunitz trypsin inhibitor, others include cobra venom factor (CVF), L-amino acid oxidase (LAOO) and phosphodiesterases (PDE)) (See also Figure S1G).

tissue was markedly lower compared to organoids, most likely due to presence of non-epithelial cells in gland tissue. Amongst the toxins, three-finger toxins (3FTx) were most abundant (**Figure 2E**). We also detected expression of cysteine-rich secretory proteins (CRISP), snake venom metalloproteinases (SVMP) and kunitz type protease inhibitors (KUN) (**Figure 2E and S1G**). The relative abundance of these toxin classes matched the venom gland tissue at transcriptome level, and the crude venom composition at protein level (Whiteley et al., 2019) (**Figure 2E**). Of note, the seven-day differentiation protocol increased overall toxin gene expression but reduced the expression of CRISP (**Figure 2D and E**). We concluded that *A. l. cowlesi* organoids produced a near-normal spectrum of venom factors.

Adult *Naja nivea* venom gland-derived organoids

To further demonstrate the long-term propagation capacity of adult venom gland derived cells using our protocol, we expanded organoids from *Naja nivea*, the Cape cobra. Venom glands from a euthanized adult individual (>1-year-old) were dissociated and cultured using the same conditions as used above (**Figure 3A**). Organoids recapitulated the epithelial phenotype and were expanded for over 18 passages (**Figure 3A and B**). While cells were viable and proliferating in 'expansion medium', we noticed reduced swelling (smaller lumen) in *N. nivea* organoids making it more difficult to mechanically split these cultures. Upon additional activation of cyclic AMP using forskolin (FSK), organoids exhibited improved swelling allowing easier splitting (**Figure 3C**). For further passaging (after passage 5) of *N. nivea* organoids, we supplemented expansion medium with FSK. Forskolin-induced swelling is well known in primary human intestinal organoids, where it is used to monitor the ability to transport chloride ions (Dekkers et al., 2013).

Exposure to 'differentiation medium', including the withdrawal of FSK resulted in the expected phenotype of less proliferative organoids (**Figure 3D**) accompanied by increased expression of *3FTx* and *KUN* by qPCR (**Figure 3E**). To establish a 'deep' gene expression profile, we performed bulk mRNA sequencing of *N. nivea* organoids in differentiation medium. PolyA enriched reads could be mapped to the *de novo A. l. cowlesi* transcriptome (**Figure 2C**). We thus identified a number of putative toxins enriched in organoids from *N. nivea*, such as beta-bungarotoxin-like, kappa-bungarotoxin-like and sarafotoxin-like (**Figure 3F**). Three-finger toxins are the main venom components of elapid snakes. Utilizing the high sequence conservation of *3FTX*, we were able to PCR-amplify the coding sequence of several variants starting from *A. l. cowlesi* and *N. nivea* organoid cDNA (**Figure 3G**). When translating these *3FTX* coding sequences *in silico* we detected peptide sequences that have not been described before in the NCBI database and are specific to the individual species (**Figure 3H**).

As our access to *A. l. cowlesi* venom gland material was more regular (one clutch a year), we focused our further characterization on organoid lines derived from *A. l. cowlesi*.

Organoids display cellular heterogeneity in toxin expression

The cellular heterogeneity of the venom gland epithelium has largely been described morphologically. We have recently demonstrated the usefulness of organoids for the detailed
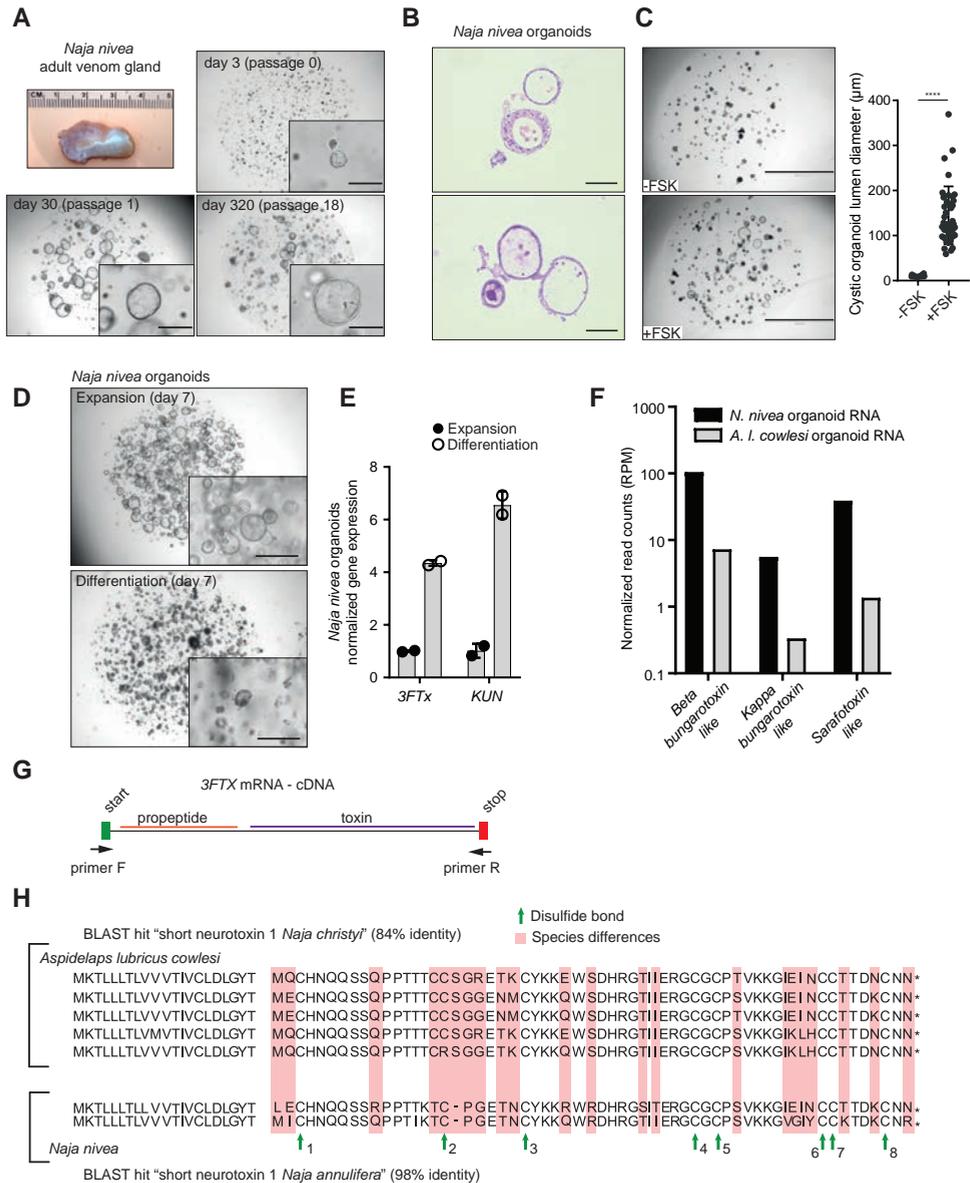
Figure 3. *Naja nivea* **organoids derived from adult venom gland reveal species-specific toxins. (A)** Adult *Naja nivea* venom gland (n=2) and organoid outgrowth after seeding of primary cells (passage 0) until passage 18 (Related to Figure S1A). Scale bars, 200 μm. **(B)** HE and PAS staining of *Naja nivea* organoids derived from an adult venom gland. Scale bars, 50 μm. **(C)** Brightfield images and quantification of *Naja nivea* organoids grown in complete expansion medium with or without supplementation of forskolin (FSK) and matching quantification. Data points represent individual organoids. Scale bars, 2000 um. **(D)** Brightfield images of *Naja nivea* organoids after 7 day expansion or differentiation protocol. Scale bars, 200 μm. **(E)** Gene expression of toxins (*3FTX* and *KUN*) in *N. nivea* organoids upon exposure to expansion or differentiation medium (7 days). Determined by qPCR, normalized to *ACTB* and relative to expansion medium. Data points represent biological replicates. **(F)** Plot of normalized read count of three selected toxins for *N. nivea* and *A. l. cowlesi* organoid bulk mRNA sequencing after 7 days of differentiation. *N.* ▸

delineation of cell lineages in the enteroendocrine compartment of the gut (Beumer et al., 2018). Using a similar strategy, we performed single cell RNA sequencing of organoids in expansion and differentiation medium and compared it to their primary tissue counterparts obtained from *A. l. cowlesi* late embryonic venom gland. From organoids, a total of 1,536 cells were sorted, processed using the SORT-seq method (Muraro et al., 2016) and analyzed using the RaceID3 package (Herman et al., 2018). Reads were mapped to the *de novo* assembled *A. l. cowlesi* transcriptome, processed using a newly generated pipeline and filtered for >2,000 transcripts per cell. The 1,092 cells which passed the thresholds displayed a median expression of 10,480 transcript counts per cell (**Figure S2A and S2B**). K-medoids-based clustering compartmentalized the cells into 12 different cell clusters, as visualized by *t*-distributed stochastic neighbor embedding (t-SNE) (**Figure 4A**). Cells derived from expansion and differentiation medium clustered mostly separately (**Figure 4B**). Expression of one of the most abundant three-finger toxins (*Pr-SNTX*) (**Figure S1G**) revealed the presence of at least 4 cell clusters producing venom factors (**Figure 4C**). In line with the bulk transcriptome data, the vast majority of venom-producing cells were derived from differentiated organoids.

Single cell RNA sequencing of freshly isolated venom gland tissue (using the same pipeline as for the organoid cells) yielded 1,255 cells which passed the same threshold (**Figure 4D**). Based on their transcriptomic profile, these cells fell into 20 different cell clusters (**Figure 4D**). Using mammalian markers of cell types expected to be present in glandular organs, we determined the following composition of our dataset: 53% epithelial cells (*EPCAM, KRT8*) (**Figure 4E and S2C**), 27% stromal cells (*COL3A1*), 8% hematopoietic cells (*HEMGN, LYZ*), 7% smooth muscle cells (*ACTA2*) and 4% endothelium (*CDH5*) (**Figure S2C**). Expression of three-finger toxin *Pr-SNTX* was highest in two of the cell clusters of epithelial origin (**Figure 4F**). We also detected strong co-expression of protein disulfide isomerase (*PDI*) with 3FTx variants (*Scutellatus toxin 3*) in the organoids (**Figure 4G**) as well as in the primary tissue cells (**Figure 4H**). This enzyme is a key factor to ensure correct disulfide bond folding (Wang and Tsou 1993), conceivably supporting the disulfide bond-rich structure of three finger toxins. A more detailed analysis of toxin-related gene expression per cluster uncovered that individual venom factors were strongly enriched in separate organoid clusters, suggesting the presence of specialized cells for some of the toxin families (**Figure S2D**).

We the extracted all 670 epithelial cells from the dataset and performed reclustering (**Figure S2E**). This resulted in 10 different epithelial subclusters, each enriched for different venom factors. Comparing the heterogeneity in the expression of different toxin classes (*3FTx, CRISP, KUN, SVMP and CTL*) we concluded that organoids successfully recapitulate the cellular complexity of venom producing cells *in vivo* (**Figure S2F**).

▸ *nivea* data generated from bulk RNA sequencing mapped on de *novo A. l. cowlesi* transcriptome (n=1). *A. l cowlesi* read count from differentiation data Figure 3 (n=1). (**G**) Schematic representation of PCR amplification of *3FTX* coding sequence from cDNA obtained from organoids. Location of forward (F) and reverse (R) primer. (**H**) Peptide sequence of *in silico* translated *3FTX* coding sequences of *A. l. cowlesi* and *N. nivea.* Highlighted in red species differences, green arrow conserved disulfide sites. BLASTP hit for best match with proteins in the NCBI database.
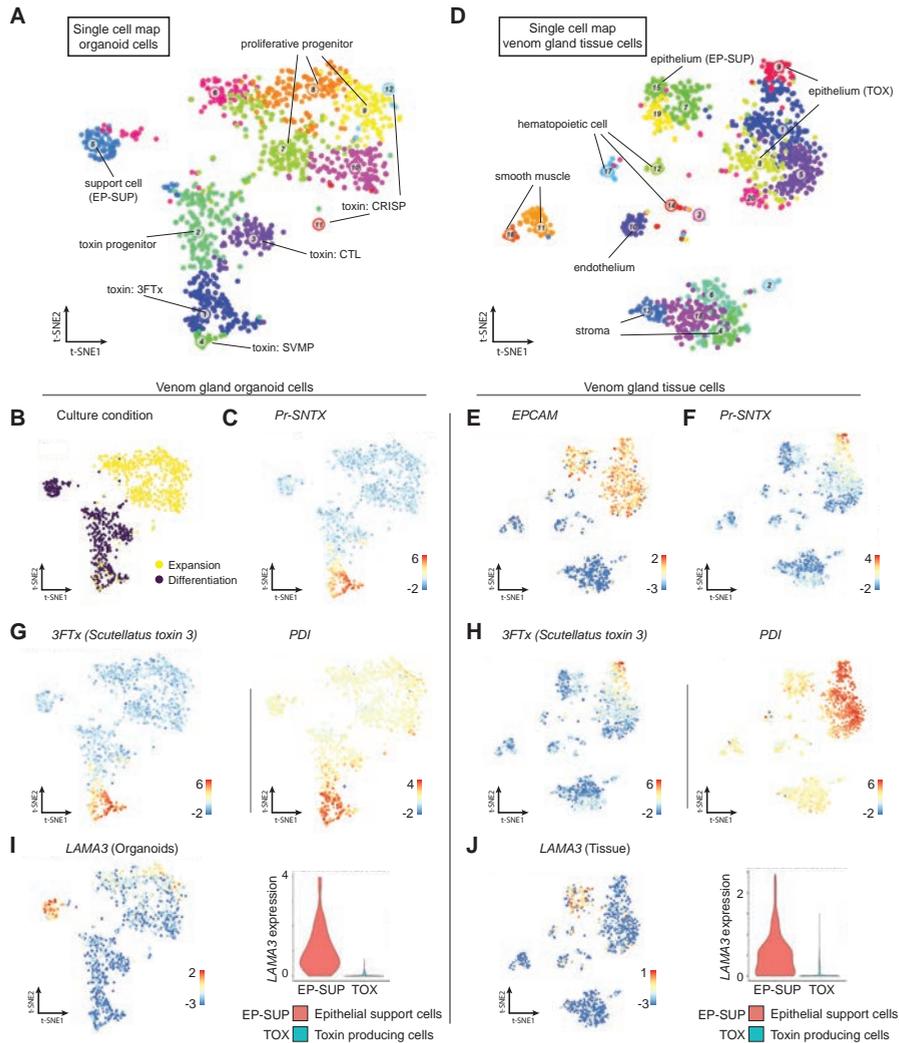
Figure 4. Single cell transcriptome analysis of organoids and primary tissue reveals distinct venom gland cell types, see also Figure S2, S3, S4 and S5. (A) Single cell RNA sequencing: clustering of *A. l. cowlesi* venom gland organoid cells (n=1092) visualized by *t*-distributed stochastic neighbor-embedding (*t*-SNE) map. Colors highlight different clusters (n=12). **(B)** *t*-SNE map indicating exposure to expansion medium (yellow) or differentiation medium (purple) for seven days. **(C)** Expression level of three-finger toxin *Pr-SNTX* in *t*-SNE map (color coded logarithmic scale of transcript expression). **(D)** Clustering of primary *A. l. cowlesi* venom gland tissue cells (n=1255) visualized by *t*-distributed stochastic neighbor-embedding (*t*-SNE) map. Colors highlight different clusters (n=20). **(E)** Expression level of epithelial cell marker *EPCAM* in *t*-SNE map (color coded logarithmic scale of transcript expression). **(F)** Expression level of three-finger toxin *Pr-SNTX* in *t*-SNE map (color coded logarithmic scale of transcript expression). **(G-H)** Expression levels of selected genes in *t*-SNE map (color coded logarithmic scale of transcript expression). Left is venom gland organoids cells and right is primary venom gland tissue cells. **(I-J)** Expression levels of *LAMA3* in *t*-SNE map (color coded logarithmic scale of transcript expression) and violin plots visualizing expression levels of cluster-enriched toxins. For both organoid cells (I) and venom gland tissue cells (J) color coded EP-SUP cells (red) and TOX cells (green).

Cluster 1 in organoids and cluster 6 in primary tissue comprised cells enriched for transcripts of *3FTX* genes (**Figure S3A**), a subset of these cells additionally expressed Kunitz variants (*KUN*) (**Figure S3B**). The cells in cluster 4 (organoids) and cluster 6 and 8 (tissue) were positive for SVMP (**Figure S3C**). Organoid cells expressing CRISP genes were enriched in cluster 11 and 12; in the tissue these genes were found to be expressed in a larger number of cells enriched in cluster 3 (**Figure S3D**). Organoid cluster 3 consisted of cells co-expressing C-type lectin (CTL) and Waprin-related toxins (**Figure S3E**) (Torres et al., 2003; Ogawa et al., 2005). CTLs have not previously been detected in the *Aspidelaps lubricus* venom proteome and did not form an independent cluster in the tissue dataset (Whiteley et al., 2019).

Cluster 5, containing cells exclusively derived from differentiated organoids, was devoid of any known toxin expression. Comparing this cluster with venom producing clusters 1, 2, 3 and 4, we found these cells to be enriched in extracellular matrix component transcripts such as laminin (*LAMA3*) (**Figure 4I**). This transcriptomic separation is indicative of two different cellular lineages captured by the organoids, an "epithelial supportive cell" fate (EP-SUP) and a "toxin producing cell" fate (TOX) (**Figure S4A**). In the complete dataset of venom gland tissue cells, we found cluster 7, 15 and 19 to be enriched in *LAMA3* expression, while these cells expressed much fewer toxin transcripts compared to the other epithelial cells (**Figure 4J and S4B**). In organoids as well as tissue, the EP-SUP lineage additionally expressed *CTGF*, *COL7A1* and *FRZB* (**Figure S4C-E**). Based on our single cell sequencing of the venom gland tissue, all epithelial lineage/cell types were represented in the organoids. Importantly, marker expression analysis of each of the 12 cell clusters in early and late passage organoids supported the notion that cellular composition of organoids was stable over time in culture (**Figure S4F**).

Characterization of the Wnt-activated proliferative cells

Having identified the distribution of toxin expression in differentiated cells, we next focused on the expression patterns specific to cell clusters observed under expansion conditions. While some cells in expansion medium clustered together with toxin-expressing cells in cluster 1, 4, 11 and 12, the vast majority fell within the expansion medium-specific clusters 6–10 (**Figure 4A and B**). Cluster 10 was distinct in displaying high expression of phospholipase A2 inhibitor (*PLI*) (**Figure S5A**), previously described in other snake species as a self-protective mechanism against venom PLA2 toxins (Lima et al., 2011). Expression of *PLI* was also found in epithelial cells from primary tissue (**Figure S5B**).

Cells in organoid clusters 7, 8 and 9 expressed the proliferation marker Ki-67 (*MKI67*) (**Figure 5A**). Transcripts for the snake homologs of the mammalian stem cell markers *RNF43* (Koo et al., 2012)*, ASCL2* (van der Flier et al., 2009) and *LGR5* (Barker et al., 2007) were enriched in these clusters and were specifically observed under expansion conditions, while being rare in primary tissue (**Figure 5A, 5B and S5C - F**). This implied these clusters to represent reptilian adult stem/progenitor cells.

Addition of R-spondin 3 enhanced expression of the stem cell markers *LGR5*, *ASCL2* and *RNF43* together with the well-established Wnt targets *TCF7L2* and *AXIN2*. Supplementation of exogenous Wnt3a further increased expression of these marker genes. As expected, stimulation
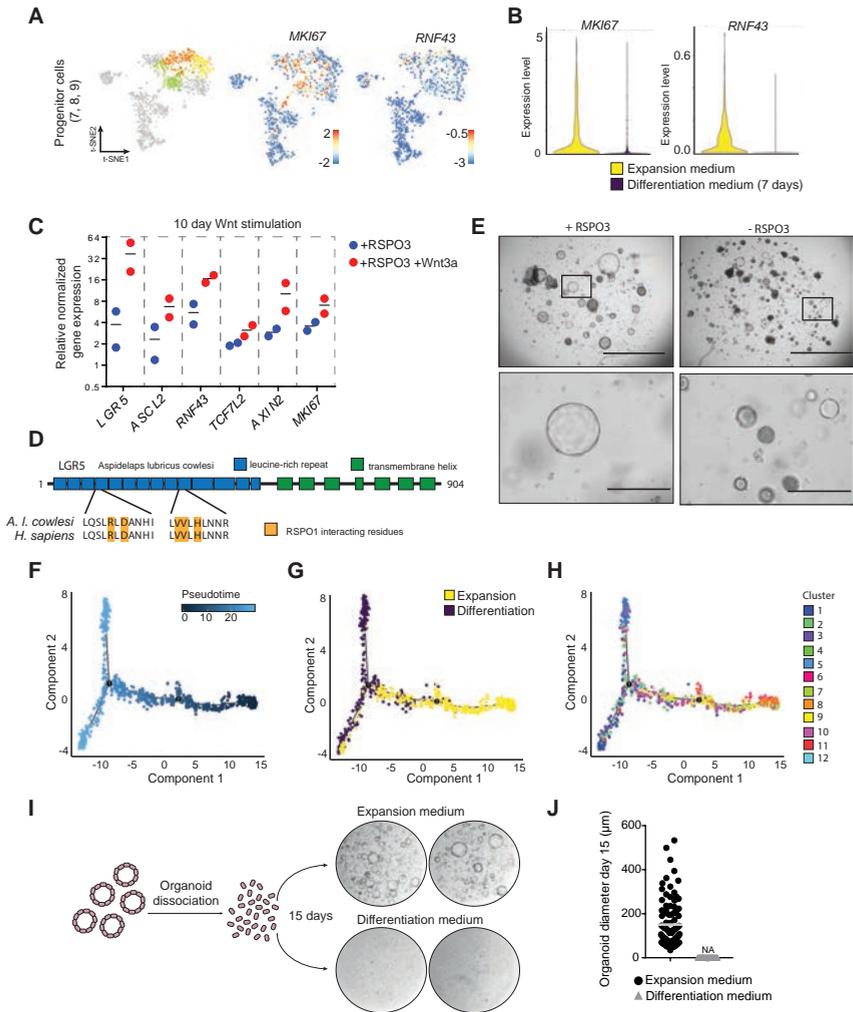
Figure 5. Non-venomous organoid cells include Wnt-active proliferating cells, see also Figure S5.
(**A**) Expression levels of *MKI67* and *RNF43* in *t*-SNE map (color coded logarithmic scale of transcript expression). (**B**) Violin plot of *MKI67* and *RNF43* expression levels in cells from expansion medium (yellow) or differentiation medium (purple). (**C**) Changes in organoid gene expression levels after 10 day Wnt activation (addition of RSPO and RSPO plus exogenous Wnt3a). Expression levels were determined by qPCR and shown relative to 'no RSPO and no Wnt3a exposure', normalized to *ACTB*. (**D**) Schematic overview of *A. l. cowlesi* LGR5 protein and alignment of R-spondin-interacting residues with the human amino acid sequence. (**E**) Representative brightfield images of organoids grown for 14 days with (+RSPO3) or without (-RSPO3) R-spondin in culture medium. Scalebars, 2000 μm (upper panels) and 400 μm (lower panels). (**F**) Ordering of cells from single cell sequencing data along a pseudotemporal trajectory using Monocle. (**G**) Position of cells exposed to expansion medium (yellow) or differentiation medium (purple) for seven days along the pseudotemporal trajectory. (**H**) Position of cells belonging to the 12 different clusters along the pseudotemporal trajectory. (**I-J**) Schematic overview of experimental setup and representative brightfield images of organoid outgrowth in expansion medium and differentiation medium starting from a near-single cell state. Circular images represent 20 μL BME droplets containing cells/organoids. J is quantification of I. Data points represent individual organoids.

of the Wnt pathway also induced the proliferation marker *MKI67* (**Figure 5C**). *A. l. cowlesi* LGR5 shares the same leucine-rich repeats and transmembrane helices with high conservation in the R-spondin interacting residues (R144, D146, V213, V214 and H216) compared to the human protein (**Figure 5D**) (Chen et al., 2013). Indeed, human R-spondin 3 was found to be essential for organoid expansion (**Figure 5E, 1D and S1B**).

The non-epithelial niche is replaced in organoid culture by the addition of defined exogenous growth factors. Using the single cell sequencing dataset of venom gland tissue, we searched for expression of such secreted factors. We detected the EP-SUP cluster as an epithelial source of *WTN10A* expression (**Figure S5G**). Stromal cells (cluster 4, 6, 13 and 18) specifically expressed *WNT9A*, *FGF7* and the BMP-antagonists *CHRD* (Piccolo et al., 1996) and *FSTL1* (Sylva et al., 2011) (**Figure S5H - K**). As expected, transcripts for these proteins with a stromal source *in vivo* were not detected in the organoids.

Pseudotemporal ordering of venom gland organoid cells grown in expansion- and differentiation medium using Monocle (Trapnell et al., 2014) was then applied to shed light on the venom gland stem cell hierarchy. Cells from the proliferative progenitor clusters 7, 8 and 9 were placed at the start of the pseudotime axis, in agreement with their stem/progenitor state. A bifurcation was formed by the EP-SUP cells of cluster 5 and toxin-producing (TOX) cells from clusters 1, 2, 3 and 4 (**Figure 5F - H**). To confirm the lack of proliferative and stem cell capacity in the mature clusters (cluster 1 - 5, **Figure 4A and B**), we differentiated organoids using the growth factor depletion described above. This caused a dramatic reduction in outgrowth and proliferative capacity compared to expansion conditions (**Figure 5I and J**).

Regional heterogeneity in toxin production in the venom gland is maintained in organoid culture

Secretory cells of other mammalian organs, such as the intestine, display regional variation in the types of secreted products. For instance, organoids derived from different regions of the mouse small intestine produce a region-specific repertoire of hormones that is maintained indefinitely *in vitro* (Beumer et al., 2018). We investigated regional heterogeneity in toxin production in the snake venom gland using our organoid culture system. The venom gland of embryonic *A. l. cowlesi* was dissected into a proximal (located near the duct) and a distal part and established region-specific organoids (**Figure 6A**). After culturing organoids for 1 month (4 passages), we analyzed toxin expression in expansion and differentiation medium by qPCR. We found *CTL* expression to be strongly enriched in 'proximal' organoids, whereas 'distal' organoids cells predominantly produced *3FTX* and *KUN* toxins (**Figure 6B**).

Next, we utilized RNA *in situ* hybridization to visualize the expression of these toxins in venom gland tissue. We found a strong enrichment of *CTL* transcripts in the proximal part of the gland, whereas *KUN* was predominantly expressed in the distal tissue, confirming our findings in organoids (**Figure 6C, S5L and M**). This is in line with a previous report of *CTL* expression in the proximally located accessory gland in the king cobra (Vonk et al., 2013). *CRISP* expression was homogenous along the proximal-distal axis of the gland, while displaying a bias towards basal over luminal cells (**Figure 6C**). These data highlighted regional heterogeneity
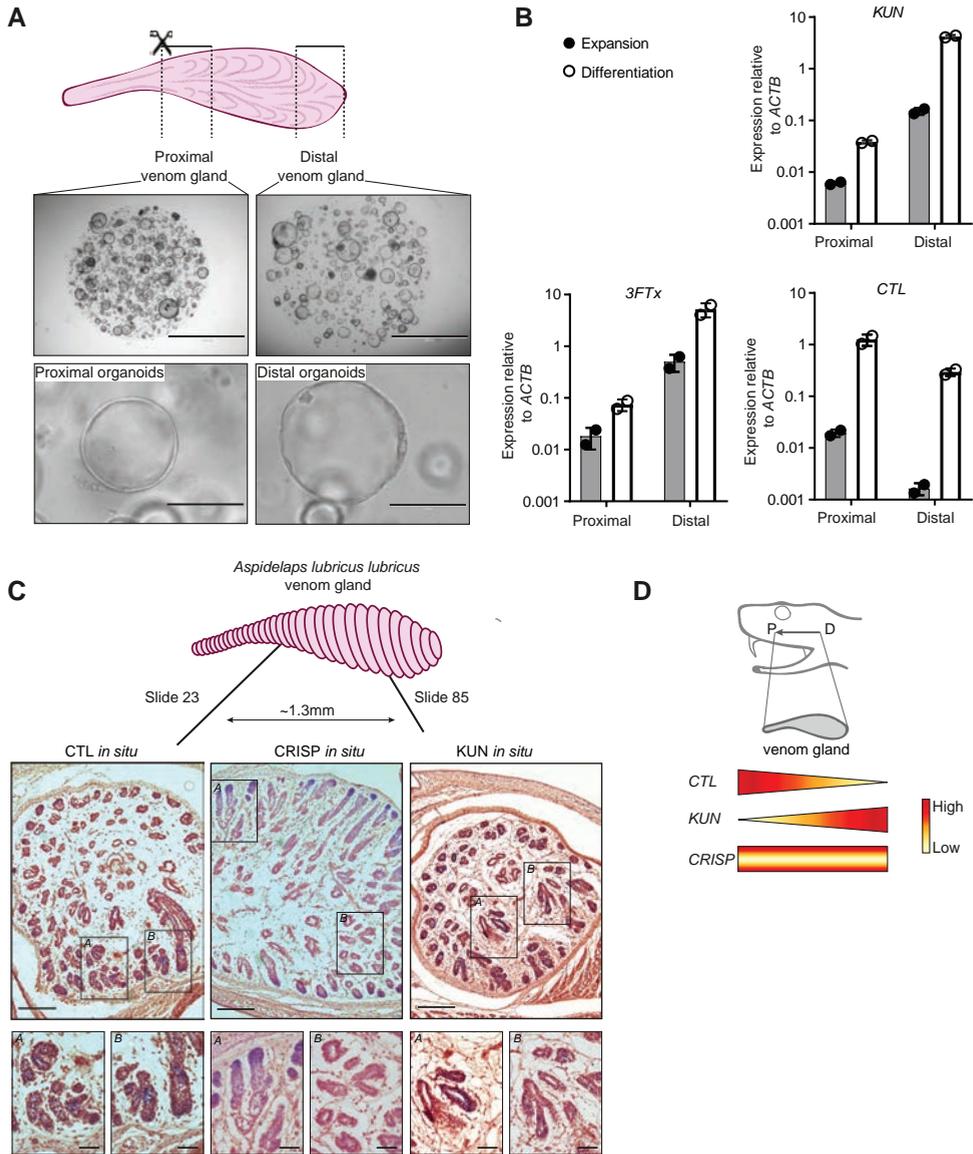
Figure 6. Hard-wired regional heterogeneity in toxin expression is maintained in organoids from proximal and distal venom gland, see also Figure S5. (A) Schematic representation and brightfield images of organoids derived from the proximal and distal region of the *A. l. cowlesi* venom gland. Scale bars, upper panels 2000 μm, lower panels 200 μm. **(B)** Gene expression of toxins (*KUN, 3FTX and CTL*) in proximal and distal organoid lines (passage 4) exposed to expansion medium or differentiation medium (7 days). Expression levels were determined by qPCR shown relative to *ACTB*. Data points represent biological replicates. **(C)** Schematic representation of venom gland sections from proximal to distal. And BM-purple stain of in situ hybridization for CTL (slide 23), CRISP and KUN (slide 85) in *A. l. cowlesi* venom gland tissue. Scale bars, upper panels 200 μm, lower panels 50 μm. **(D)** Schematic overview of regional heterogeneity in venom gland toxin expression: CTL expression proximal, KUN expression distal and CRISP expression along the proximal-distal axis on the basal side.

in toxin expression in the snake venom gland (**Figure 6D**). Long-term maintenance of this phenotype in organoid culture showed this not to be determined by extrinsic (non-epithelial) growth factors.

Venom gland organoids secrete functionally active toxins

To investigate whether the stem cell-based organoids produced functionally active venom components, we analyzed organoid protein extracts after 7 days of differentiation. The presence of secretory vesicles and the apparent accumulation of proteins in the lumen of organoids suggested significant production of secretory proteins (**Figure 7A and 2B**). To validate functional translation and secretion of toxins, we generated a fluorescent three-finger toxin reporter organoid line. Using CRISPR-HOT (Artegiani et al., unpublished data), we tagged one of the endogenous three-finger toxins (*NTX4*) with a fluorescent protein (mNeon-Green) (**Figure S6A**). We detected the green fluorescent fusion protein inside the cells of the organoids and accumulating in the lumen (**Figure S6A**).

We also directly compared organoid extract (**Figure 7B**) with *A. l. cowlesi* crude venom using liquid chromatography-mass spectrometry (LC-MS). This revealed three main peaks between 6-8 kDa of near identical mass between the two samples (**Figure 7C**). Tryptic digest analysis of these peptides yielded patterns compatible with several venom-related proteins such as 3FTx and CRISP (**Figure S6B**).

Snake venom neurotoxins are described to act primarily on acetylcholine receptors (nAChR and mAChR) (Nirthanan et al., 2004; Karlsson et al., 2000). As an initial assessment of the biological activity of organoid-secreted venom peptides, we exposed murine muscle cultures (endogenously expressing the nAChR) to the organoid supernatant and recorded signal propagation using calcium imaging in an OrganoPlate (Trietsch et al., 2013) (**Figure 7D, S6C - F**). This organ-on-a-chip platform allows live cell imaging during venom exposure in 96 independent mature muscle cultures per plate. The venom peptides derived from differentiated organoids abolished the stimulatory effect of the acetylcholine receptor agonist carbachol on muscle cells, as did the positive control alpha-bungarotoxin (α-BTX). Supernatant from non-venom producing (human) organoids did not inhibit calcium wave propagation after addition of carbachol (**Figure 7E, S6G - I**).

In a similar set-up, we exposed rat cortical neurons grown on microelectrode array (MEA) plates to organoid supernatant and recorded neuronal activity (**Figure S7A**). Upon acute exposure, we noticed a marked increase in neuronal activity, following a similar trend observed with recombinant alpha-bungarotoxin (**Figure S7B - D**). This occurred most likely through an antagonistic effect on inhibitory GABAa-receptors (McCann et al., 2006), without affecting neuronal cell viability (**Figure S7E - G**).

DISCUSSION

We report the establishment of reptilian adult stem cell-based organoids. Using a mammalian niche growth factor cocktail, these cells can be expanded seemingly indefinitely. The ability
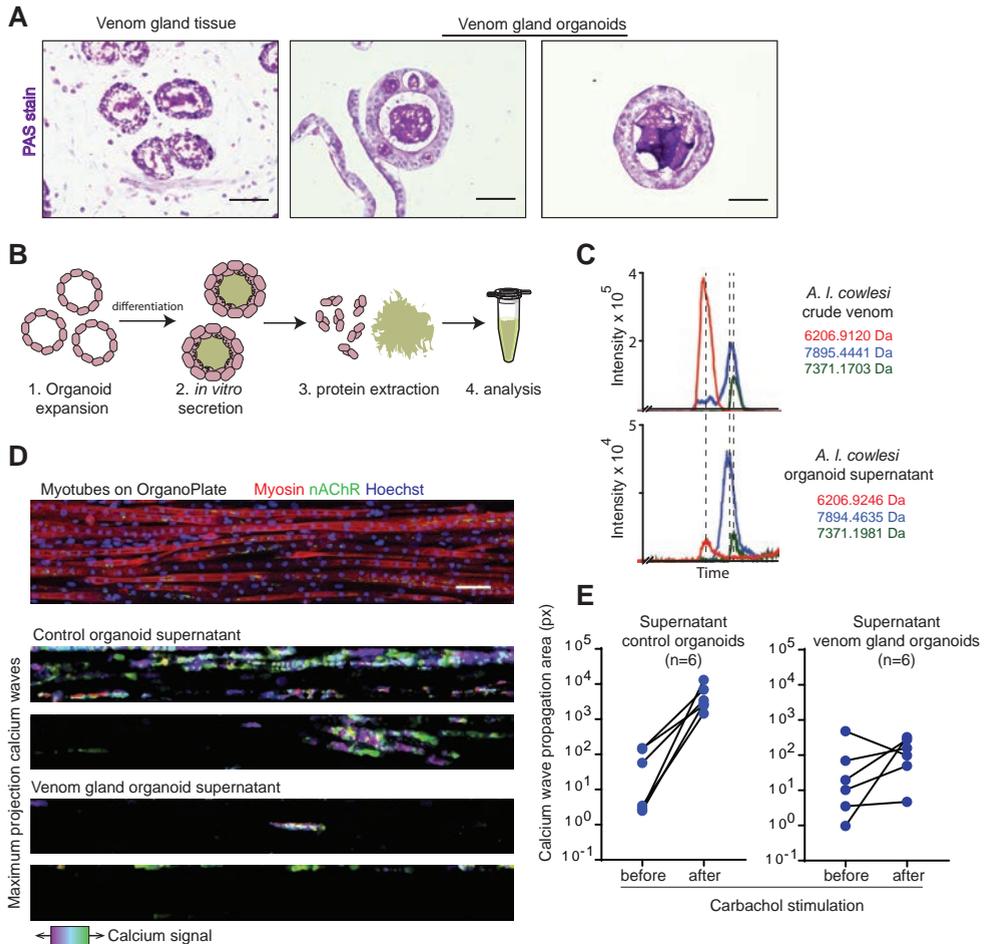
Figure 7. Organoids produce biologically active toxins, related to Figure S6 and S7. (A) Luminal accumulation of secreted proteins was determined using periodic acid-Schiff (PAS) staining in late embryonic venom gland tissue and organoids. Scale bars, 50 μm. **(B)** Schematic representation of organoid extraction. Organoids were expanded for four days and then subjected to a seven-day differentiation protocol. In PBS, the lumen was opened using mechanical shearing and cells disrupted using sonication. Cellular debris was pelleted by centrifugation and supernatant collected for analysis. **(C)** Direct comparison of *A. l. cowlesi* crude venom (upper panel) with organoid extract (lower panel) using liquid chromatography–mass spectrometry (LC-MS) identifies three overlapping products. Ion-chromatogram displays signal intensity (y-axis) over retention time in minutes (x-axis) with indicated mass per trace. **(D)** Immunofluorescence staining of C2C12 myotubes in OrganoPlate, myosin (red), nAChR (green) and Hoechst (blue) (upper panel). Scale bar, 100 μm. The calcium wave propagation was imaged before and after exposure to different supernatants in chips. Maximum projected images depict the total calcium signal over all image frames after carbachol stimulation, in which purple reflects activity to the left and green to the right (lower panels). **(E)** Quantification of (D). The total calcium signal was calculated before and after carbachol stimulation based on total fluorescent signal. The venom gland organoid supernatant blunts the induction of calcium sparks. Data points represent biological replicates.

to culture venom gland epithelial cells over long periods of time represents a platform for a comprehensive understanding of the biology of the venom gland and the various venom constituents. Short-term cultures of different Viperidae snakes have previously been reported, such as explants of *Bothrops jararaca* (Carneiro et al., 2006; Yamanouye et al., 2006), suspension cultures of *Bitis gabonica* (Sells et al., 1989) and unpolarized two-dimensional cell lines of *Crotalus durissus terrificus* (Duarte et al., 1999). The venom gland organoid cultures provides advantages in structural conservation (**Figure 1C**), cellular heterogeneity (**Figure 4**), regional heterogeneity (**Figure 6**), long-term expansion (**Figure 1B and 3A**), genetic modifiability (**Figure S1E and S6A**) and broad applicability across snake species (**Figure S1A**). With a *de novo*-generated transcriptome of *A. l. cowlesi,* we demonstrate that the venom gland organoid cells produce a diverse spectrum of venom factors over many passages (**Figure 2**).

Previous mammalian organoid cultures have highlighted the importance of Wnt-driven stem cell maintenance *in vitro* (Clevers 2016). Indeed, we identify a proliferative population of cells in the venom gland organoids that is defined by the expression of the snake homologs of established Wnt-driven stem cell genes of mammalian epithelia: *LGR5 (*Barker et al., 2007), ASCL2 (van der Flier et al., 2009) and RNF43 (Koo et al., 2012). The importance of R-spondin in reptilian cell proliferation highlights the critical role and a high level of conservation of the Wnt pathway in ASC biology. We believe that the culture conditions used here may be widely applicable to vertebrate species.

While a divergent composition of snake venoms between species, and even between and within individuals, is well-established (Casewell et al., 2014; Augusto-de-Oliveira et al., 2016; Zancolli et al., 2019), the cellular heterogeneity underlying the production of individual toxins is not well described. Initial attempts at characterizing the cell types of the snake venom gland through electron microscopy has resulted in the morphological identification of at least 4 cell types, of which columnar epithelial cells containing granules have been suggested to be the main source of venom production and secretion (Mackessy 1991). The amount and type of venom detectable in individual cells is influenced by the secretory cycle, referring to temporal dynamics in venom production and secretion (Taylor et al., 1986; Shaham and Kochva 1969). A previous study identified heterogeneity in the binding of single-component antibodies to secretory granules and attributed this mainly to the secretory cycle-dependent production of toxins rather than cellular heterogeneity of the venom gland (Taylor et al., 1986).

Our data describe cellular heterogeneity in the venom gland epithelium at the level of transcription. Gene expression profiling at single cell resolution of both organoids and primary tissue uncovers the coincident expression of characteristic groups of toxins by individual cell types. Future studies may use venom gland organoids to dissect the stimuli and timing of venom production and secretion. The ability to indefinitely expand these organoids and repeatedly harvest venom supernatants in a highly defined environment may help overcome hurdles posed by the significant variation in snake venom composition.

Finally, the current study opens new avenues for bioprospecting of snake venom components and may be developed into a production platform for (modified) snake venom, allowing novel therapeutic strategies to tackle snakebite.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

Y.P., J.P., J.B. and H.C. conceptualized the project, designed the experiments, interpreted the results and wrote the manuscript. H.B. and J.K performed immunohistochemistry experiments. J.P., B.D.B., A.V.O., M.C.H. and H.J.G.S. performed bulk and single cell mRNA sequencing analysis on *de novo* transcriptome. Y.E.B.E. assisted with FACS experiments. C.L.I., W.J.V.D.W. and P.J.P. performed transmission electron microscopy. M.A.G.D.B. performed *in situ* hybridization experiments. R.L.V.I. and A.C.R. generated immunofluorescent images. J.S. and J.K. performed mass spectrometry analysis. S.A., T.D.K. and N.R.C. performed toxin gene expression analysis on bulk RNA sequencing data. N.R.W., X.M.S. and T.O. performed and analyzed toxicity tests on Mimetas OrganoPlate. H.M.H., F.J.V. and M.K.R. assisted in initializing the project and provided access to venom gland tissue. R.H.S.W. and R.G.D.M.K performed and analyzed toxicity tests on rat cortical neurons in MEA plates. W.G. and N.R.C. provided venom gland tissue.

DECLARATION OF INTERESTS

H.C. is inventor on several patents related to organoid technology; his full disclosure is given at https://www.uu.nl/staff/JCClevers/. N.R.W. and T.O. are employees of MIMETAS BV, the Netherlands, which is marketing the OrganoPlate. OrganoPlate is a registered trademark of MIMETAS.

STAR METHODS
Lead contact and materials availability

Further information and requests for resources and reagents should be directed to the Lead Contact, Hans Clevers (h.clevers@hubrecht.eu).

Unique/stable reagents generated in this study are available and can be requested from the Lead Contact, a completed Materials Transfer Agreement may be required.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Snakes

All animal procedures complied with local ethical guidelines. Adult *Crotalus atrox* (n=2) venom glands were purchased from Natural Toxins Research Center in Texas, USA. Adult *Echis ocellatus* (n=1) and *Deinagkistrodon acutus* (n=1) venom glands were obtained from snakes maintained at the Liverpool School of Tropical Medicine, UK. Adult *Naja pallida* (n=1), *Naja nivea* (n=1) and *Bitis arietans* (n=2) venom glands were obtained from residual post-mortem material from Serpo, Rijswijk, The Netherlands. Embryonic *Naja atra* (n=2), *Naja annulifera* (n=3) and *Aspidelaps lubricus cowlesi* (n=7, two different nests) from eggs were obtained from local breeders in The Netherlands. All snakes were captive bred and maintained in individual cages within a temperature, humidity and light-controlled environment according to local protocols for husbandry of venomous snakes. Sex of the animals was not determined as it was presumed to have no biological impact on the experimental procedure or conclusion of the study. Non-invasive venom extraction ('milking') was performed with adult *Echis ocellatus, Deinagkistrodon acutus, Naja pallida*, *Naja nivea* and *Bitis arietans* during lifetime prior to euthanizing and venom gland isolation for organoid culture.

Venom gland organoid cultures

Cells from a single gland were plated in approximately 100 µL Cultrex Pathclear Reduced Growth Factor Basement Membrane Extract (BME) (3533-001, Amsbio). After BME solidification, culture medium was added. Expansion culture medium was based on AdDMEM/F12 (Gibco) supplemented with B27, Glutamax, HEPES, 100 U/mL Penicillin-Streptomycin (all Thermo-Fisher), 100 mg/mL Primocin (Invivogen), 1.25mM N-acetylcysteine, 10 mM Nicotinamide (both Sigma-Aldrich) and the following growth factors: 2% Noggin conditioned medium (U-Protein Express), 2% Rspo3 conditioned medium (U-Protein Express), 50 ng/mL EGF (Peprotech), 0.5 µM A83-01 (Tocris), 1 µM PGE2 (Tocris), 100 ng/mL FGF10 (Peprotech), 100 nM Gastrin (Tocris). Naja nivea organoid cultures were supplemented with 1 µM FSK (Tocris). For the first seven days after seeding, the expansion medium was supplemented with 10 µM Y-27632 (Abmole). In case of bacterial contaminations following antibiotics were added: 50 mg/mL Gentamicin (Sigma), 2.5 µg/ml Ciprofloxacin (Sigma-Aldrich), 20 µM Erythromycin (Sigma-Aldrich) and 100 nM Azithromycin (Sigma-Aldrich). Ten days after seeding organoids were removed from the BME, mechanically dissociated into small fragments using a Pasteur pipette and re-seeded in fresh BME. Passage was performed in 1:3 – 1:5 split ratio once every ten days for at least 6 months.

C2C12 cell line

The C2C12 mouse myoblast cell line was purchased from Sigma-Aldrich (91031101). Sigma is partnered with the European Collection of Authenticated Cell Cultures (ECACC) repository. Cells were cultured in T75 flasks (734-2705, Corning) in DMEM (11965092, Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS, 16140-071, Thermo Fisher

Scientific), 2 mM glutamine (G7513, Sigma-Aldrich) and 1% penicillin/streptomycin (P4333, Sigma-Aldrich) at 37°C. C2C12 cells were routinely tested for mycoplasma contamination and were found negative.

Primary rat cortical neurons

Primary rat cortical cells were isolated from postnatal day 0-1 pups of timed pregnant Wistar rat dams (Envigo) as described previously (Dingemans et al., 2016). Briefly, pups were decapitated and cortices were rapidly dissected on ice and kept in dissection medium (Neurobasal®-A supplemented with 25 g/L sucrose, 450 µM L-glutamine, 30 µM glutamate, 1% penicillin/streptomycin and 10% FBS, pH 7.4) during the entire procedure. Cortices were dissociated to a single-cell suspension by mincing with scissors, trituration and filtering through a 100 µm mesh (EASYstrainer, Greiner). The cell suspension was diluted to a 2 x 106 cells/mL solution. Droplets of 50 µL were placed on the electrode fields in wells of 48-wells MEA plates (Axion BioSystems) pre-coated with 0.1% PEI solution diluted in borate buffer (24 mM sodium borate/50 mM boric acid in Milli-Q, pH adjusted to 8.4). Cells were left to adhere for 2 hours before adding 450 µL dissection medium. Primary rat cortical cultures were kept at 37°C in a humidified 5% CO2 incubator. At day *in vitro* (DIV) 2, 90% of the dissection medium was replaced with glutamate medium (Neurobasal®-A supplemented with 25 g/L sucrose, 450 µM L-glutamine, 30 µM glutamate, 1% penicillin/streptomycin and 2% B27 supplement, pH 7.4) to prevent glial overgrowth. At DIV4, 90% of the glutamate medium was replaced with FBS medium (Neurobasal®-A supplemented with 25 g/L sucrose, 450 µM L-glutamine, 1% penicillin/ streptomycin and 10% FBS, pH 7.4).

METHOD DETAILS
Venom gland isolation

In the last week of development (~day 53 – 58), snakes were removed from the egg by an incision in the shell. To avoid bacterial contamination, eggs were briefly rinsed with bleach and 70% ethanol prior to opening in a laminar flow cabinet. Animals were directly euthanized by decapitation. The late-embryonic and adult venom glands were surgically removed and stored on ice in AdDMEM/F12 (Gibco) supplemented with penicillin/streptomycin.

Venom gland organoid cultures

Isolated venom glands were chopped into small pieces of approximately 1 mm using a scalpel. Muscle and connective tissue were removed and discarded as much as possible. Epithelial tissue pieces were enzymatically digested in 5 mL collagenase (Sigma-Aldrich, C9407, 1 mg/mL) with 10 µM ROCK inhibitor Y-27632 (Abmole, M1817) in AdDMEM/F12 (Gibco) for about 30 minutes while shaking (120 RPM) at 32 °C. The homogenous cell suspension was pelleted and washed twice with AdDMEM/F12 prior to plating.

For differentiation experiments, established organoids were passaged and grown in expansion medium for four days. After four days, the medium was removed and residual

medium washed away two times with PBS, after which differentiation medium was added containing; a base of AdDMEM/F12 (Gibco) supplemented with GlutaMAX, HEPES, 100 U/mL Penicillin-Streptomycin (all Thermo-Fisher), 100 µg/mL Primocin (InvivoGen), 1.25mM N-acetylcysteine, 10 mM Nicotinamide (both Sigma-Aldrich) and 1 µM PGE2 (Tocris) applied for seven days.

Quantification of cell viability/outgrowth efficiency was performed using CellTiter-Glo® accoring to the manufacturer's protocol.

Wnt-conditioned medium was produced as previously described, and used at 50% of the medium volume (Sato et al., 2011).

Images of organoid cultures were taken on EVOS FL Cell Imaging System (Thermo Fisher). Video of beating cilia in organoid was captured (20 frames/second, 20x objective) on EVOS FL Auto 2 (Thermo Fisher).

pLV lentiviral vector containing a CMV promoter driving the expression of TagRFP and Puro resistance gene (via IRES) was transduced in venom gland organoids.

Immunohistochemistry and imaging

Organoids were harvested in cell recovery solution (354253, Corning) and fixed in 4% formaldehyde solution (Sigma-Aldrich) for at least 2 hours at room temperature. Venom glands used for immunohistochemistry (n=3) were directly embedded in 4% paraformaldehyde upon dissection and fixed for at least 2 hours. Samples were washed and dehydrated by an increasing ethanol gradient before embedding in paraffin. Sections were cut and hydrated before staining. Hematoxylin and eosin (H&E) and periodic acid–Schiff (PAS) staining on organoids and tissue was performed as previously described (Sato et al., 2009). Slides were imaged using a Leica DM4000 microscope.

3D imaging of organoids was performed as described previously (Dekkers et al, 2019). In short, organoids were harvested using ice cold PBS and fixed in 4% paraformaldehyde at 4°C for 45 min. Organoids were then washed with PBT (PBS, 0.1% Tween) and incubated overnight with primary antibodies, B-catenin (H-102, Santa Cruz) and b-tubulin (H-235 Santa Cruz). The next day, organoids were washed with PBT (PBS, 0.1% Tween) and incubated overnight at 4 °C with secondary antibodies, Alexa-fluor-647 Phalloidin (both from Thermo Fisher Scientific) and DAPI (Invitrogen). Dense organoids were optically cleared overnight in a glycerol-fructose clearing solution prior to imaging. Cystic organoids were not optically cleared to prevent these from collapsing. Organoid imaging was performed on a Zeiss LSM 880 using a 10x dry and 25x oil immersion objective. Imaris imaging software was used for 3D rendering of images.

For immunofluorescence in the OrganoPlate, cultures were fixed and stained as previously described (Wevers et al., 2018). The following antibodies were used: anti-myosin (A4.1025, DSHB), anti-desmin (ab8470, Abcam), anti-dystrophin (ab15277, Abcam), goat anti-mouse AlexaFluor 555 (A21422, Thermo Fisher Scientific), goat anti-mouse AlexaFluor 647 (A21236, Thermo Fisher Scientific), goat anti-rabbit AlexaFluor 555 (A32732, Thermo Fisher Scientific) and donkey anti-rabbit 647 (SAB4600177, Sigma-Aldrich). The nicotinic acetylcholine receptor was visualized using α-bungarotoxin AlexaFluor 488 (B13422, Thermo Fisher Scientific). Nuclei

were stained using Hoechst (H3570, Thermo Fisher Scientific). All steps were performed at room temperature (RT). Cells were imaged using the Micro XLS-C HCI System (Molecular Devices). A 3D reconstruction of a myotube culture was made using Fiji.

In situ hybridization

Late embryonic A. l. cowlesi heads (n=2) were fixed in 4% PFA overnight at 4°C. Fixed heads were divided in two through the midline of the body, separating the left and right venom gland. Material was dehydrated using sequential methanol steps (25, 50, 75, 100%), washed with three times for one hour 100% ethanol and embedded in paraffin. Sections of 10 μm were deparaffinized (nucleoclear, ethanol, Milli-Q) and pretreated (PBST, ProtK 10 min, PBS and 4% PFA). For hybridization, slides were covered with hybridization mix for one hour at 60°C (2% blocking powder, 50% formamide, 5x SSC, 1mg/mL tRNA, 50 μg/mL heparin, 0,1% Triton X-100, 0,1% CHAPS and 5mM EDTA) prior to overnight incubation with probe mix (800 - 1000 ng/mL) at 60°C. On the next day, the slides were washed (2x SSC, 0,1% CHAPS, 50% Formamide) for 30 minutes three times prior to overnight anti-Dig-AP (in 10% sheep serum) incubation at 4°C. Signal was visualized using BM-purple and counterstained with 0,1% neutral red. Images were taken on a Leica DM4000 microscope.

Following primers were used for probe generation:

CRISP	F_TGCTGCAACAGTCTTCTGGAAC	R_ATATAGTTTGCATGAAGGGCATCA
CTL	F_TCTGGGGATTCTGCCTCTTG	R_ACTTGCAGATGAAGGGCAGG
KUN	F_CCCTGCTTAACTTCCCCCAA	R_GGCAGGGTCTCCAGGAAGG

Electron microscopic analysis

Aspidelaps lubricus organoid lines (n=2 organoid lines derived from different individuals) were harvested in cell recovery solution (354253, Corning) to remove BME and fixed by chemical fixation using 3% glutaraldehyde in cacodylate buffer. The fixation was followed by 1% osmium tetroxide plus 1.5% potassium ferrocyanide postfixation. Then, the organoids were dehydrated in series of ethanol and gradually embedded in Epon resin. Ultrathin sections were observed in a Tecnai Spirit T12 Electron Microscope equipped with an Eagle CCD camera (Thermo Fisher Scientific).

Organoid protein extraction

Aspidelaps lubricus organoids were grown and differentiated as described above. For LC-MS analysis, a minimum of 500 μL of organoid containing BME (~25 droplets) was harvested for protein extraction. In short, differentiation medium was removed and any residual liquid removed with PBS washes, organoids were collected and mechanically dissociated in 500 μL PBS using a glass Pasteur pipette in a 1.5 mL eppendorf tube. Tubes were sonicated for 15 cycles (30 sec on, 30 sec off) using the Bioruptor Plus (Diagnode). Any leftover debris was spun down at 12,000 rpm, the supernatant collected, snap frozen on dry ice and stored at -80 °C.

For functional toxicity readout, a minimum of 500 µL of organoid containing BME was washed in the tissue culture plate with PBS and supplemented with C2C12 differentiation medium for two days prior to collection of supernatants. The same procedure was applied to harvest supernatant from human colon control organoid cultures as a negative control (Sato et al., 2011).

Liquid chromatography–mass spectrometry and Mascot database search

For comparison of *A. l. cowlesi* crude venom and organoid extracts, liquid chromatographic separation was performed with parallel at-line nanofractionation and mass spectrometry analysis. MS data of venom and organoid samples was analyzed for identical masses found in the samples to confirm production of venom toxins by organoids.

LC separation was performed using a Shimadzu UPLC system ('s Hertogenbosch, The Netherlands), and a 250×4.6 mm Waters Xbridge Peptide BEH300 C18 analytical column with a 3,5-µm particle size combined with a 300-Å pore size. Separations were performed in a Shimadzu CTD-30A column oven at 30°C. Mobile phase A comprised of 98% H2O, 2% ACN and 0.1% FA, and mobile phase B comprised of 98% ACN, 2% H2O and 0.1% FA. A linear increase of mobile phase B from 0% to 50% in 20 min was followed by a linear increase from 50% to 90% B in 4 min and a 5 min isocratic separation at 90% B. The starting conditions were reached again in 1 min and the column was then equilibrated for 10 min at 0% B. A post-column flow split in 1:9 ratio directed the larger fraction to a FractioMateTM FRM100 nanofraction collector (SPARK-Holland & VU, Netherlands, Emmen & Amsterdam) or a modified Gilson 235P autosampler. To further confirm that venom toxins were present in the organoid extracts, crude organoid extracts were subjected to tryptic digestion after which these were analyzed with nanoLC-MS/MS. For this, 20 µL of each organoid sample was transferred to an Eppendorf tube containing 15 µl of digestion buffer (25 mM NH4HCO3, pH 8.2). Subsequently, 1.5 µl of reducing agent (0.5% β-mercaptoethanol) was added to each tube followed by a 10-min incubation step at 95°C. Post incubation, the samples were cooled to room temperature and centrifuged at 150 RCF for 10 seconds in a Himac CT 15RE centrifuge. Next, 3 µL of alkylating agent (55 mM iodoacetamide) was added to the tubes before incubation in the dark at room temperature for 30 min. Next, 3 µL of 0.1 µg/µL trypsin was added before incubation at 37°C for 3 h, and then an additional 3 µl of trypsin was added to the tubes and the samples were incubated overnight. Subsequently, 1 µL of 5% FA was added to quench the digestion, followed by brief centrifugation to remove any particulate matter. Finally, the samples were centrifuged for 10 s at 150 RCF and the supernatants were transferred to autosampler vials with glass inserts and analyzed with nanoLC-MS/MS. The data files obtained were run through the Mascot search engine (MASCOT; Matrix Science, London, United Kingdom) against the Swiss-Prot, NCBInr database.

Myotube cultures

C2C12 cells were harvested and resuspended to obtain a cell suspension of 5,000 cells/µL in cold medium supplemented with 1:25 Matrigel-GFR (356231, Corning). The C2C12 cell suspension was seeded in the perfusion channel of an OrganoPlate 2-lane (9603-400B, Mimetas) against a collagen-I ECM gel as previously described (Wevers et al., 2018). The plate was incubated on a flat surface for 3 hours to allow cell attachment, after which medium perfusion was started with C2C12 differentiation medium composed of DMEM supplemented with 2% horse serum (26050-088, Thermo Fisher Scientific), 1 µM insulin (I9278, Sigma-Aldrich), 2 mM glutamine, and 1% penicillin/streptomycin. From day 2-5, the differentiation medium was supplemented with 10 µM cytarabine (C1768, Sigma-Aldrich) to inhibit cell proliferation.

For cell viability assay, C2C12 cells were exposed for 30 minutes to supernatant from *Aspidelaps lubricus* venom gland organoids, supernatant from human colon organoids, recombinant alpha bungarotoxin and a killing control. After that exposure was terminated and fresh medium added for 24h and 48h. WST8 assay was used according to the manufacturer's protocol to quantify cell viability.

Calcium imaging

Calcium imaging assays were performed at day 8 after plating. C2C12 myotubes in the OrganoPlate were incubated with 20 µM Cal-520 (ab171868, Abcam), 0.04% Pluronic-127 (P6866, Thermo Fisher Scientific) in C2C12 differentiation medium without serum for 60 minutes at 37°C, followed by 30 more minutes at RT. Next, cultures were exposed to C2C12 differentiation medium (negative control), human colon organoid supernatant (negative control), snake venom gland organoid supernatant or 10 µM α-bungarotoxin (positive control) for 30 minutes. Calcium imaging recordings were made using an ImageXpress Micro XLS-C Confocal High-Content Imaging System (Molecular Devices, wide field mode, 4x magnification, 2 Hz). An individual chip was imaged for 90 seconds to record baseline calcium activity. Myotubes in the chip were then exposed to 500 µM carbachol or vehicle (0.5% water) and immediately placed back in the microscope for another 90 seconds of calcium imaging to record the response to carbachol or vehicle. This pipeline was repeated for each chip.

The calcium imaging recordings were processed using scripts developed in Fiji (Schindelin et al., 2012). Recordings were corrected for camera-induced noise via Kalman Filtering and corrected for bleaching. Movement of fluorescent signal was quantified via Gaussian Window MSE resulting in RGB-encoded videos showing movement of fluorescent signal to the left and right side in purple and green, respectively. The colored areas were measured per timeframe and averaged over time. All videos were processed to correct for flow-induced artefacts caused by movement of fluorescent cell debris. In the case of calcium wave propagation, movement was detected to a similar extent in both the left and right direction (<2 fold difference) and the average area of the two directions was determined. In the case of flow of fluorescent debris, the movement detected in one direction is dominant over the other (>=2 fold difference) and movement in the dominant direction was excluded.

MEA recordings rat cortical neurons

Spontaneous electrical activity was recorded on the day of exposure (DIV9-11) as described previously (Nicolas et al., 2014; Tukker et al., 2018). In short, signals were recorded using a Maestro 768-channel amplifier with integrated heating system and temperature controller and a data acquisition interface (Axion BioSystems). Data acquisition was managed with Axion's Integrated Studio (AxIS 2.4.2.13) and recorded as .RAW files. All channels were sampled at the same time with a gain of 1200x and a sampling frequency of 12.5kHz/channel with a 200-500 Hz band-pass filter. Prior to the recording, MEA plates were allowed to equilibrate for ~10 min in the Maestro.

Spontaneous neuronal activity was recorded prior to exposure to generate a baseline recording. Immediately following this recording, cells were exposed to the test compounds (α-bungarotoxin, organoid supernatant and vehicle control) and activity was recorded for another 30 min. Each well was exposed to only one single concentration of one compound in order to prevent receptor (de)sensitization.

To determine (modulation of) spontaneous activity, .RAW data files were re-recorded to obtain Alpha Map files. In this re-recording, spikes were detected with the AxIS spike detector (Adaptive threshold crossing, Ada BandFlt v2) and a variable threshold spike detector set at 7x standard deviation (SD) of internal noise level (rms) on each electrode. Post/pre-spike duration was set to 3.6/2.4 ms respectively. For further data analysis, spike files were loaded in NeuralMetric Tool (version 2.2.4, Axion BioSystems) and only active electrodes (MSR ≥ 0.1 spikes/s) in active wells (≥ 1 active electrode) were included in data analysis. The effects of test compounds on spontaneous activity were determined by comparing the baseline activity with activity following exposure. A custom-made MS Excel macro was used to calculate treatment ratios (TR) per well for the mean spike rate, which were normalized to appropriate vehicle control.

For cell viability, rat cortical cells were cultured at a density of 3.0 x 10^4 cells/well. At DIV9-10, cells were exposed for 24 h to the test compounds, after which cell viability was assessed using a combined CFDA-AM / neutral red / alamar blue assay. In short, cells were incubated for 30 minutes with 12.5 mM alamar blue and 4 mM CFDA-AM in FBS at 37 °C. Resorufin was measured spectrophotometrically at 540/590 nm (Infinite M200 microplate; Tecan), whereas hydrolyzed CFDA was measured spectrophotometrically at 493/541 nm. The aB/CFDA solution was then replaced by 200 mL neutral red solution (175 mM in PBS, Invitrogen) for 1 h at 37 °C. Next, cells were incubated for 30 min with 200 mL extraction solution (1% glacial acetic acid, 50% ethanol, and 49% H2O) during gentle shaking at room temperature. After 30 min extraction, fluorescence was measured spectrophotometrically at 530/645 nm.

RNA extraction and *de-novo* transcriptome assembly

RNA isolations of organoids and tissues for *de-novo* transcriptome assembly and bulk RNA sequencing were performed using TRIzol (Invitrogen) following the manufacturer's instructions. Organoid RNA isolations for quantitative PCR were performed with RNeasy

Mini Kit (Qiagen) following the manufacturer's instructions. Quantitative PCR analysis was performed using the SYBR Green and Bio-Rad systems. Changes in expression were calculated using CFX manager software (Bio-Rad). Primers were designed using NCBI primer design tool with the coding sequence of annotated scaffolds of interest as input. Primers used in this study are;

ACTB, F_CTGGCCTAGGACACAGTACG,R_GCTCAGACTCCATTGCAACA

*LGR5,*F_GTTCCCCTTCCTGCATGTCT,R_ACCAAACTAGCATCTTTTGCCTT

*ASCL2,*F_CACTCGGCTTATTCGTCGGA,R_CTCCCGAACCAACTGGTGAA

*AXIN2,*F_GATAGAAGCTGAGGCAGCCC,R_CCCCTTCGCATGTCCTCTAC

*RNF43,*F_TTCCCATGAGTTCCATCGGC,R_GGCGGTACCTGATGTTGACT

*TCF7L2,*F_GCTATCACCGGGCACTGTAG,R_GGTCCTCACGAGATTGCCTG

*MKI67,*F_CAGGTGCATGAATCTGGTATTGAA,R_ATTTAGCGCTGCTTCTGTGACC

*3FTX,*F_GTGGTGGTGACAATCGTGTG,R_GGTTGCGATGACTGTTGGTT

*KUN,*F_GTCCAGGACTCTGTGAACTGC,R_GCATTGTTTTGCAGCCAGGTT

*CTL,*F_TACACCCCAGGAACCCTTCT,R_TATTGGTGACGGAGACGCAC

*HSPA8,*F_AGCAGTACAAAGCGGAGGAC,R_TCTGCCGTGCTCTTCATGTT

Sequencing and de-novo transcriptome assembly was performed by BaseClear B.V. as follows: PolyA enriched RNA from liver tissue, pancreas tissue, venom gland tissue, organoids early passage, organoids late passage expansion and organoids late passage differentiation was sequenced. Paired-end sequence reads were generated using the Illumina HiSeq2500 system. FASTQ sequence files were generated using bcl2fastq2 version 2.18. Initial quality assessment was based on data passing the Illumina Chastity filtering. Subsequently, reads containing PhiX control signal were removed using an in-house filtering protocol. The second quality assessment was based on the remaining reads using the FASTQC quality control tool version 0.11.5. The transcriptome assembly was performed with Trinity (version 2.4.0, Haas et al., 2013), using the combined data of all six samples. The resulting consensus assembly was filtered to remove all sequences shorter than 500bp, resulting in 311,948 remaining scaffolds. These were aligned with BLASTN (version 2.6.0) against a local copy of the NCBI-NT database (built in February 2018). Toxin-encoding contigs were validated by manual sequence analysis (e.g. ORF identification) of those exhibiting annotations consistent with known snake venom toxin proteins.

Bulk and single cell RNA sequencing

Bulk RNA sequencing of *N. nivea* was performed by the Utrecht Sequencing Facility (uSeq) as follows: Libraries were prepared based on polyA enrichment. Paired-end sequence reads were generated using the NextSeq500 system.

For organoid single cell RNA sequencing, *A. l. cowlesi* organoids were grown and differentiated as described above. BME was removed by washing in ice-cold DMEM (Gibco). For tissue sequencing, embryonic *A. l. cowlesi* venom glands were dissociated with collagenase I as described above and subsequently processed in the same way as organoids. Organoids or tissue

fragments were resuspended in TrypLE Express (Gibco) pre-heated to 32°C and dissociated under repeated pipetting. Upon reaching single cell state, samples were pelleted, washed, resuspended in FACS buffer (advanced DMEM, 10 mM Y-27632 and DAPI) and strained (35 μm). Cells were immediately sorted into 384-well plates containing ERCC spike-ins (Agilent), RT primers and dNTP (Promega) using a BD FACSJazz (BD Biosciences). Plates were prepared using Mosquito HTS (TTPlabtech). scRNA-seq libraries were prepared following the SORT-seq protocol (Muraro et al., 2016), which is based on the CEL-seq2 method (Hashimshony et al., 2016). Briefly, cells were first lysed for 5 min at 65°C, and RT and second-strand mixes were dispensed by the Nanodrop II liquid handling platform (GC Biotech). Single-cell double-stranded cDNAs were pooled together and *in vitro* transcribed for linear amplification. Illumina sequencing libraries were prepared using the TruSeq small RNA primers (Illumina) and these DNA libraries were sequenced paired-end at 60 and 26 bp read length, respectively, on the Illumina NextSeq.

Bulk and single cell RNA sequencing analysis

Bulk paired-end reads of *A. l. cowlesi* and *N. nivea* tissue and organoids were mapped to the *de novo* assembled *A. l. cowlesi* transcriptome using STAR. When a read mapped to multiple scaffolds, a fractional count of 1/x was recorded, where x is the number of scaffolds the read maps to. To reduce the complexity of the count table, scaffolds with identical blast hits were merged by summation of their transcript counts. All scripts used for read counting are deposited at https://github.com/BuysDB/reptilianOrganoids. For single cell analysis, paired-end reads were demultiplexed on the cell barcode and UMI and adapters trimmed using Cutadapt 2. Next, the reads were mapped and counted as described for the bulk samples, additionally we removed amplification duplicates by UMI-scaffold combinations. Scaffold annotations were filtered to exclude the terms ("mitochondrial", "mitochondrion", "ribosomal", "rRNA", "ribosomal", "microsatellite", "transposon", "SINE", "LINE repeat", "noBlast") and requiring them to be detected in at least 2 cells, resulting in 42,507 unique annotations. Subsequently, cells were filtered by requiring at least 2000 counts and at least 100 different genes to be expressed, leaving 1,092 cells (1,255 for tissue) to be included for dimensionality reduction and k-medoids-based clustering using RaceID3 according to the vignette (Herman et al., 2018), starting with an expected cluster number of 10 (organoid) and 20 (tissue). Violin plots were generated using Seurat functions (Butler et al. 2018) (version 3.0.0.9000) on the dataset analyzed by RaceID3. Single cell trajectories were analyzed using Monocle (version 2.6.4) (Trapnell et al., 2014). In short, genes were selected as input features based on differential expression between expansion and differentiation medium and dimensionality was reduced for display using reverse graph embedding.

QUANTIFICATION AND STATISTICAL ANALYSIS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to the sample allocation during experiments

and outcome assessment. All data are presented as mean ± standard error of the mean (SEM), unless stated otherwise. Value of n is always displayed in the figure as individual data points, more information can be found in the figure legends. Statistical tests included unpaired two-tailed t-test for Figures 1D, S1C, S6I, S7E/F/G and paired two-tailed t-test for Figure S7B/C/D using GraphPad Prism. ns p > 0.05, * p ≤ 0.05 **, p ≤ 0.01 ***, p ≤ 0.001, **** p ≤ 0.0001.

DATA AND CODE AVAILABILITY

All bulk and single cell RNA sequencing data of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE129581.

SUPPLEMENTAL VIDEOS (ON CELL.COM)

Video S1. Three-dimensional imaging of *Aspidelaps lubricus cowlesi* organoids, related to Figure 1

Part 1; DAPI (orange), Beta-catenin (yellow), Tubulin (red) and Actin (blue).
Part 2; DAPI (orange), Beta-catenin (red), Tubulin (green) and Actin (blue).

Video S2. Functionally beating cilia in VGO (20x objective, 20 frames/second), related to Figure 2 and S1

REFERENCES

Ainsworth, S., Slagboom, J., Alomran, N., Pla, D., Alhamdi, Y., King, S.I., Bolton, F.M.S., Gutierrez, J.M., Vonk, F.J., Toh, C.H., et al. (2018). The paraspecific neutralisation of snake venom induced coagulopathy by antivenoms. Communications biology 1, 34.

Artegiani, B., and Clevers, H. (2018). Use and application of 3D-organoid technology. Human molecular genetics 27, R99-R107.

Augusto-de-Oliveira, C., Stuginski, D.R., Kitano, E.S., Andrade-Silva, D., Liberato, T., Fukushima, I., Serrano, S.M., and Zelanis, A. (2016). Dynamic Rearrangement in Snake Venom Gland Proteome: Insights into Bothrops jararaca Intraspecific Venom Variation. Journal of proteome research 15, 3752-3762.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., et al. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449, 1003-1007.

Beumer, J., Artegiani, B., Post, Y., Reimann, F., Gribble, F., Nguyen, T.N., Zeng, H., Van den Born, M., Van Es, J.H., and Clevers, H. (2018). Enteroendocrine cells switch hormone expression along the crypt-to-villus BMP signalling gradient. Nature cell biology 20, 909-916.

Bolhaqueiro, A.C.F., van Jaarsveld, R.H., Ponsioen, B., Overmeer, R.M., Snippert, H.J., and Kops, G. (2018). Live imaging of cell division in 3D stem-cell organoid cultures. Methods in cell biology 145, 91-106.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature biotechnology 36, 411-420.

Carneiro, S.M., Zablith, M.B., Kerchove, C.M., Moura-da-Silva, A.M., Quissell, D.O., Markus, R.P., and Yamanouye, N. (2006). Venom production in long-term primary culture of secretory cells of the Bothrops jararaca venom gland. Toxicon : official journal of the International Society on Toxinology 47, 87-94.

Casewell, N.R., Wagstaff, S.C., Wuster, W., Cook, D.A., Bolton, F.M., King, S.I., Pla, D., Sanz, L., Calvete, J.J., and Harrison, R.A. (2014). Medically important differences in snake venom composition are dictated by distinct postgenomic mechanisms. Proceedings of the National Academy of Sciences of the United States of America 111, 9205-9210.

Chen, P.H., Chen, X., Lin, Z., Fang, D., and He, X. (2013). The structural basis of R-spondin recognition by LGR5 and RNF43. Genes & development 27, 1345-1350.

Clark, G.C., Casewell, N.R., Elliott, C.T., Harvey, A.L., Jamieson, A.G., Strong, P.N., and Turner, A.D. (2019). Friends or Foes? Emerging Impacts of Biological Toxins. Trends in biochemical sciences.

Clevers, H. (2016). Modeling Development and Disease with Organoids. Cell 165, 1586-1597.

de Lau, W., Barker, N., Low, T.Y., Koo, B.K., Li, V.S., Teunissen, H., Kujala, P., Haegebarth, A., Peters, P.J., van de Wetering, M., et al. (2011). Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. Nature 476, 293-297.

Dekkers, J.F., Alieva, M., Wellens, L.M., Ariese, H.C.R., Jamieson, P.R., Vonk, A.M., Amatngalim, G.D., Hu, H., Oost, K.C., Snippert, H.J.G., et al. (2019). High-resolution 3D imaging of fixed and cleared organoids. Nature protocols 14, 1756-1771.

Dekkers, J.F., Wiegerinck, C.L., de Jonge, H.R., Bronsveld, I., Janssens, H.M., de Winter-de Groot, K.M., Brandsma, A.M., de Jong, N.W., Bijvelds, M.J., Scholte, B.J., et al. (2013). A functional CFTR assay using primary cystic fibrosis intestinal organoids. Nat Med 19, 939-945.

Dingemans, M.M., Schutte, M.G., Wiersma, D.M., de Groot, A., van Kleef, R.G., Wijnolts, F.M., and Westerink, R.H. (2016). Chronic 14-day exposure to insecticides or methylmercury modulates neuronal activity in primary rat cortical cultures. Neurotoxicology 57, 194-202.

Duarte, M.M., Montes de Oca, H., Diniz, C.R., and Fortes-Dias, C.L. (1999). Primary culture of venom gland cells from the South American rattlesnake (Crotalus durissus terrificus). Toxicon : official journal of the International Society on Toxinology 37, 1673-1682.

Fatehullah, A., Tan, S.H., and Barker, N. (2016). Organoids as an in vitro model of human development and disease. Nature cell biology 18, 246-254.

Fox, J.W., and Serrano, S.M. (2008). Exploring snake venom proteomes: multifaceted analyses for complex toxin mixtures. Proteomics 8, 909-920.

Fry, B.G., Vidal, N., van der Weerd, L., Kochva, E., and Renjifo, C. (2009). Evolution and diversification of the Toxicofera reptile venom system. Journal of proteomics 72, 127-136.

Gehart, H., van Es, J.H., Hamer, K., Beumer, J., Kretzschmar, K., Dekkers, J.F., Rios, A., and Clevers, H. (2019). Identification of Enteroendocrine Regulators by Real-Time Single-Cell Differentiation Mapping. Cell 176, 1158-1173 e1116.

Grun, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature 525, 251-255.

Gutierrez, J.M., Calvete, J.J., Habib, A.G., Harrison, R.A., Williams, D.J., and Warrell, D.A. (2017). Snakebite envenoming. Nature reviews Disease primers 3, 17063.

Gutierrez, J.M., and Lomonte, B. (2013). Phospholipases A2: unveiling the secrets of a functionally versatile group of snake venom toxins. Toxicon : official journal of the International Society on Toxinology 62, 27-39.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature protocols 8, 1494-1512.

Harvey, A.L. (2001). Twenty years of dendrotoxins. Toxicon : official journal of the International Society on Toxinology 39, 15-26.

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K.J., Rozenblatt-Rosen, O., et al. (2016). CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome biology 17, 77.

Herman, J.S., Sagar, and Grun, D. (2018). FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. Nature methods 15, 379-386.

Izidoro, L.F., Sobrinho, J.C., Mendes, M.M., Costa, T.R., Grabner, A.N., Rodrigues, V.M., da Silva, S.L., Zanchi, F.B., Zuliani, J.P., Fernandes, C.F., et al. (2014). Snake venom L-amino acid oxidases: trends in pharmacology and biochemistry. Biomed Res Int 2014, 196754.

Karlsson, E., Jolkkonen, M., Mulugeta, E., Onali, P., and Adem, A. (2000). Snake toxins with high selectivity for subtypes of muscarinic acetylcholine receptors. Biochimie 82, 793-806.

Kochva, E. (1987). The origin of snakes and evolution of the venom apparatus. Toxicon : official journal of the International Society on Toxinology 25, 65-106.

Koo, B.K., Spit, M., Jordens, I., Low, T.Y., Stange, D.E., van de Wetering, M., van Es, J.H., Mohammed, S., Heck, A.J., Maurice, M.M., et al. (2012). Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. Nature 488, 665-669.

Lima, R.M., Estevao-Costa, M.I., Junqueira-de-Azevedo, I.L., Ho, P.L., Diniz, M.R., and Fortes-Dias, C.L. (2011). Phospholipase A2 inhibitors (betaPLIs) are encoded in the venom glands of Lachesis muta (Crotalinae, Viperidae) snakes. Toxicon : official journal of the International Society on Toxinology 57, 172-175.

Mackessy, S.P. (1991). Morphology and ultrastructure of the venom glands of the northern pacific rattlesnake Crotalus viridis oreganus. Journal of morphology 208, 109-128.

Maimets, M., Rocchi, C., Bron, R., Pringle, S., Kuipers, J., Giepmans, B.N., Vries, R.G., Clevers, H., de Haan, G., van Os, R., et al. (2016). Long-Term In Vitro Expansion of Salivary Gland Stem Cells Driven by Wnt Signals. Stem cell reports 6, 150-162.

McCann, C.M., Bracamontes, J., Steinbach, J.H., and Sanes, J.R. (2006). The cholinergic antagonist alpha-bungarotoxin also binds and blocks a subset of GABA receptors. Proceedings of the National Academy of Sciences of the United States of America 103, 5149-5154.

Millers, E.K., Trabi, M., Masci, P.P., Lavin, M.F., de Jersey, J., and Guddat, L.W. (2009). Crystal structure of textilinin-1, a Kunitz-type serine protease inhibitor from the venom of the Australian common brown snake (Pseudonaja textilis). FEBS J 276, 3163-3175.

Muraro, M.J., Dharmadhikari, G., Grun, D., Groen, N., Dielen, T., Jansen, E., van Gurp, L., Engelse, M.A., Carlotti, F., de Koning, E.J., et al. (2016). A Single-Cell Transcriptome Atlas of the Human Pancreas. Cell systems 3, 385-394 e383.

Nicolas, J., Hendriksen, P.J., van Kleef, R.G., de Groot, A., Bovee, T.F., Rietjens, I.M., and Westerink, R.H. (2014). Detection of marine neurotoxins in food safety testing using a multielectrode array. Mol Nutr Food Res 58, 2369-2378.

Nirthanan, S., and Gwee, M.C. (2004). Three-finger alpha-neurotoxins and the nicotinic acetylcholine receptor, forty years on. Journal of pharmacological sciences 94, 1-17.

Ogawa, T., Chijiwa, T., Oda-Ueda, N., and Ohno, M. (2005). Molecular diversity and accelerated evolution of C-type lectin-like proteins from snake venom. Toxicon : official journal of the International Society on Toxinology 45, 1-14.

Piccolo, S., Sasai, Y., Lu, B., and De Robertis, E.M. (1996). Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4. Cell 86, 589-598.

Richter, K., Haslbeck, M., and Buchner, J. (2010). The heat shock response: life on the verge of death. Mol Cell 40, 253-266.

Sato, T., Stange, D.E., Ferrante, M., Vries, R.G., Van Es, J.H., Van den Brink, S., Van Houdt, W.J., Pronk, A., Van Gorp, J., Siersema, P.D., et al. (2011). Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology 141, 1762-1772.

Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., et al. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 459, 262-265.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nature methods 9, 676-682.

Sells, P.G., Hommel, M., and Theakston, R.D. (1989). Venom production in snake venom gland cells cultured in vitro. Toxicon : official journal of the International Society on Toxinology 27, 1245-1249.

Shaham, N., and Kochva, E. (1969). Localization of venom antigens in the venom gland of Vipera plaestinae using a fluorescent-antibody technique. Toxicon : official journal of the International Society on Toxinology 6, 263-268.

Slagboom, J., Kool, J., Harrison, R.A., and Casewell, N.R. (2017). Haemotoxic snake venoms: their functional activity, impact on snakebite victims and pharmaceutical promise. British journal of haematology 177, 947-959.

Sylva, M., Li, V.S., Buffing, A.A., van Es, J.H., van den Born, M., van der Velden, S., Gunst, Q., Koolstra, J.H., Moorman, A.F., Clevers, H., et al. (2011). The BMP antagonist follistatin-like 1 is required for skeletal and lung organogenesis. PLoS One 6, e22616.

Taylor, D., Iddon, D., Sells, P., Semoff, S., and Theakston, R.D. (1986). An investigation of venom secretion by the venom gland cells of the carpet viper (Echis carinatus). Toxicon : official journal of the International Society on Toxinology 24, 651-659.

Torres, A.M., Wong, H.Y., Desai, M., Moochhala, S., Kuchel, P.W., and Kini, R.M. (2003). Identification of a novel family of proteins in snake venoms. Purification and structural characterization of nawaprin from Naja nigricollis snake venom. The Journal of biological chemistry 278, 40097-40104.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nature biotechnology 32, 381-386.

Trietsch, S.J., Naumovska, E., Kurek, D., Setyawati, M.C., Vormann, M.K., Wilschut, K.J., Lanz, H.L., Nicolas, A., Ng, C.P., Joore, J., et al. (2017). Membrane-free culture and real-time barrier integrity assessment of perfused intestinal epithelium tubes. Nature communications 8, 262.

Tsetlin, V.I. (2015). Three-finger snake neurotoxins and Ly6 proteins targeting nicotinic acetylcholine receptors: pharmacological tools and endogenous modulators. Trends Pharmacol Sci 36, 109-123.

Tukker, A.M., Wijnolts, F.M.J., de Groot, A., and Westerink, R.H.S. (2018). Human iPSC-derived neuronal models for in vitro neurotoxicity assessment. Neurotoxicology 67, 215-225.

van der Flier, L.G., van Gijn, M.E., Hatzis, P., Kujala, P., Haegebarth, A., Stange, D.E., Begthel, H., van den Born, M., Guryev, V., Oving, I., et al. (2009). Transcription factor achaete scute-like 2 controls intestinal stem cell fate. Cell 136, 903-912.

Vonk, F.J., Casewell, N.R., Henkel, C.V., Heimberg, A.M., Jansen, H.J., McCleary, R.J., Kerkkamp, H.M., Vos, R.A., Guerreiro, I., Calvete, J.J., et al. (2013). The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proceedings of the National Academy of Sciences of the United States of America 110, 20651-20656.

Wang, C.C., and Tsou, C.L. (1993). Protein disulfide isomerase is both an enzyme and a chaperone. FASEB J 7, 1515-1517.

Wevers, N.R., Kasi, D.G., Gray, T., Wilschut, K.J., Smith, B., van Vught, R., Shimizu, F., Sano, Y., Kanda, T., Marsh, G., et al. (2018). A perfused human blood-brain barrier on-a-chip for high-throughput assessment of barrier function and antibody transport. Fluids and barriers of the CNS 15, 23.

Whiteley, G., Casewell, N.R., Pla, D., Quesada-Bernat, S., Logan, R.A.E., Bolton, F.M.S., Wagstaff, S.C., Gutierrez, J.M., Calvete, J.J., and Harrison, R.A. (2019). Defining the pathogenic threat of envenoming by South African shield-nosed and coral snakes (genus Aspidelaps), and revealing the likely efficacy of available antivenom. Journal of proteomics 198, 186-198.

Yamanouye, N., Kerchove, C.M., Moura-da-Silva, A.M., Carneiro, S.M., and Markus, R.P. (2006). Long-term primary culture of secretory cells of Bothrops jararaca venom gland for venom production in vitro. Nature protocols 1, 2763-2766.

Yamazaki, Y., and Morita, T. (2004). Structure and function of snake venom cysteine-rich secretory proteins. Toxicon : official journal of the International Society on Toxinology 44, 227-231.

Zancolli, G., Calvete, J.J., Cardwell, M.D., Greene, H.W., Hayes, W.K., Hegarty, M.J., Herrmann, H.W., Holycross, A.T., Lannutti, D.I., Mulley, J.F., et al. (2019). When one phenotype is not enough: divergent evolutionary trajectories govern venom variation in a widespread rattlesnake species. Proceedings Biological sciences 286, 20182735.

1

2

3

4

5

SUPPLEMENTARY FIGURES

A

Organoid outgrowth from venom gland tissue different snake species

Elapidae

Naja pallida
(Red spitting cobra)

Naja annulifera
(Snouted cobra)

Naja nivea
(Cape cobra)

Naja atra
(Chinese cobra)

Viperidae

Echis ocellatus
(West African carpet viper)

Deinagkistrodon acutus
(Pit viper)

Crotalus atrox
(Western diamondback rattlesnake)

Bitis arietans
(Puff adder)

Elapidae

Aspidelaps lubricus cowlesi
(Cape coral snake)

☐ Adult tissue venom glands
☐ Late-embryonic venom glands

B

Complete

Growth factor component
withdrawal day 14

-Rspo   -Noggin   -EGF

-A83-01   -PGE2   -FGF10

C

day 3   day 10

32°C

37°C

Cell viability day 10 (%)

****

32°C   37°C

D

○ *A. l. cowlesi*
● *N. nivea*

relative *HSPA8* mRNA
expression (2 hours)

32°C   37°C   32°C   37°C

E

Lentiviral expression histone 2B - RFP (H2B-RFP)

Brightfield   H2B-RFP   Merge

Organoid 1

Organoid 2

F

Ciliated cell in
venom gland organoid
A. l. cowlesi

G

Name	Toxin class
Oxyuranus scutellatus scutellatus toxin 3 mRNA, complete cds	3FTX
Naja sputatrix post synaptic alpha neurotoxin precursor (ntx) gene, ntx-4 allele	3FTX
Pseudechis rossignolii Pr-SNTX mRNA for three-finger toxin, complete cds	3FTX
Naja atra gene encoding cardiotoxin 4	3FTX
Laticauda laticaudata LIL gene for long chain neurotoxin, partial cds	3FTX
Ophiophagus hannah isolate OH-35 short chain alpha neurotoxin precursor	3FTX
Laticauda semifasciata (sea snake) gene for erabutoxin c	3FTX
Bungarus flaviceps muscarinic toxin-like protein mRNA, complete cds	3FTX
Naja atra clone n031_7f_1 three-finger toxin precursor, gene, partial cds	3FTX
Bungarus flaviceps non-conventional three finger toxin isoform 6 mRNA	3FTX
Pseudechis rossignolii Pr-LNTX mRNA for three-finger toxin, complete cds	3FTX
Naja kaouthia kaouthin-2 precursor, mRNA, complete cds	CRISP
Ophiophagus hannah opharin precursor, mRNA, complete cds	CRISP
Laticauda semifasciata latisemin mRNA, complete cds	CRISP
Bungarus multicinctus clone B218_4L_2 three-finger toxin precursor, genes,	Kunitz
Austrelaps labialis putative Kunitz-type proteinase inhibitor mRNA	Kunitz
Naja atra aci gene for chymotrypsin inhibitor precursor, exons 1-3	Kunitz
Naja atra mRNA for chymotrypsin inhibitor (aci gene)	Kunitz
Bungarus fasciatus L-amino acid oxidase mRNA, complete cds	LAAO
Macropisthodon rudis isolate PDE-1 phosphodiesterase gene, partial cds	PDE
Naja atra K-like metalloprotease precursor, mRNA, partial cds	SVMP
Naja naja cobrin precursor, mRNA, complete cds	SVMP
Naja atra metalloproteinase atrase A mRNA, complete cds	SVMP
Naja atra metalloproteinase atrase B mRNA, complete cds	SVMP
Drysdalia coronoides MTP4 mRNA, complete cds	SVMP
Naja atragin precursor mRNA, partial cds	SVMP
Micrurus corallinus MCOR0067C C-type lectin precursor, mRNA	CTL
Demansia vestigiata vestiginin-2 precursor, mRNA, complete cds	Waprin

◀ **Figure S1. Culture characteristics and toxins of venom gland organoids in mammalian growth factor medium, related to Figure 1 and 2. (A)** Organoid outgrowth from venom gland tissue of different Elapidae (n=5) and Viperidae (n=4) snake species. Origin of tissue was from either adult venom glands (orange) or late embryonic venom glands (blue). Scale bars, 200 μm. **(B)** Representative images of organoid outgrowth in complete medium or medium without one of the indicated components for 14 days. This dropout screen was performed on established organoids (passage 3). Scale bars, 400 μm. **(C)** Brightfield images and quantification of *A. l. cowlesi* organoids grown in complete expansion medium at 32°C or 37°C from the moment of splitting into a near single cell suspension to day 3 and day 10 after plating. Viability determined at day 10 using CellTiter-Glo, normalized to cells at 32°C. Scale bars, 2000 μm. Data points represent technical replicates. **(D)** Relative *HSPA8* expression normalized to *ACTB* determined by qPCR. For *A. l. cowlesi* and *N. nivea* organoids at 32°C and 37°C for 2 hours. Data points represent biological replicates. **(E)** Images of successful lentiviral transduction of venom gland organoids. Detection of histone 2B-RFP (H2B-RFP) stable construct integration and nuclear expression in red. Scale bars, 200 μm. **(F)** Transmission electron microscopy image of ciliated cell present in *A. l. cowlesi* organoid. Scale bar, 5 μm. **(G)** Highest expressed toxin genes in *A. l. cowlesi* venom gland organoids. Table of *de novo* scaffold BLASTN hit and toxin class.

A Transcript count per cell

B Annotations per cell

C *KRT8* — Epithelium (~53%) | *COL3A1* — Stroma (~27%) | *CDH5* — Endothelium (~4%)

HEMGN | *LYZ* — Hematopoietic cell (~8%) | *ACTA2* — Smooth muscle (~7%)

D

3FTX (scutellatus toxin 3)
3FTX (Ntx-4)
3FTX (Pr-SNTX)
3FTX (Cardiotoxin 4)
3FTX (Long chain neurotoxin)
3FTX (Alpha neurotoxin)
3FTX (Erabutoxin C)
3FTX (Muscarinic toxin)
3FTX (Three-finger toxin n031)
3FTX (Three-finger toxin isoform 6)
3FTX (Pr-LNTX)
CRISP (Kaouthin-2)
CRISP (Opharin)
CRISP (Latisemin)
Kunitz (Clone B218)
Kunitz (Chymotrypsin inhibitor 1)
Kunitz (Chymotrypsin inhibitor 2)
LAAO (L-amino acid oxidase)
PDE (PDE-1)
SVMP (K-like metalloprotease)
SVMP (Cobrin)
SVMP (Atrase A)
SVMP (Atrase B)
SVMP (MTP4)
SVMP (Atragin)
CTL (C-type lectin precursor)
Waprin (Vestiginin-2)

Expansion
Differentiation

1 3 4 11 12
Cluster

E Venom gland tissue all cells (n=1255)

reclustering

Venom gland tissue epithelial cells (n=670)

F Venom gland organoid cells | Venom gland tissue epithelial cells

3FTX (NTX4)
CRISP (Opharin)
KUN (CI)
SVMP (k-like metalloprotease)
CTL

Row Z-score

134

◄ **Figure S2. Single cell transcriptome analysis of venom gland primary tissue and organoid cells, related to Figure 4. (A)** Details on single cell sequencing analysis; transcript count per cell. **(B)** Details on single cell sequencing analysis; annotations per cell. **(C)** Single cell RNA sequencing: clustering of *A. l. cowlesi* primary venom gland tissue cells (n=1255) visualized by *t*-distributed stochastic neighbor-embedding (*t*-SNE) map. Expression level of cell type marker genes in *t*-SNE map (color coded logarithmic scale of transcript expression): epithelial cells, *KRT8* (n=670); stromal cells, *COL3A1* (n=342); hematopoietic cells, *HEMGN*, *LYZ* (n=103); smooth muscle cells, *ACTA2* (n=90); endothelial cells, *CDH5* (n=50). **(D)** Heatmap of toxin gene (y-axis) expression per individual cell (x-axis). All cells from toxin-expressing clusters 1, 3, 4, 11 and 12 are displayed for expression of most abundant toxin genes, see Figure S2B. Color coded expansion medium, differentiation medium and cluster numbers. **(E)** Schematic representation of reclustering epithelial cells (10 clusters, n=670) from total venom gland tissue dataset (20 clusters, n=1255). Visualized by *t*-distributed stochastic neighbor-embedding (*t*-SNE) map. **(F)** Heatmap of expression per toxin class (3FTx, CRISP, KUN, SVMP and CTL) of highest expressed genes (*NTX4, Opharin, Chymotrypsin inhibitor, K-like metalloprotease* and *CTL*) (y-axis) per cell (x-axis) for both organoids and epithelial tissue cells.
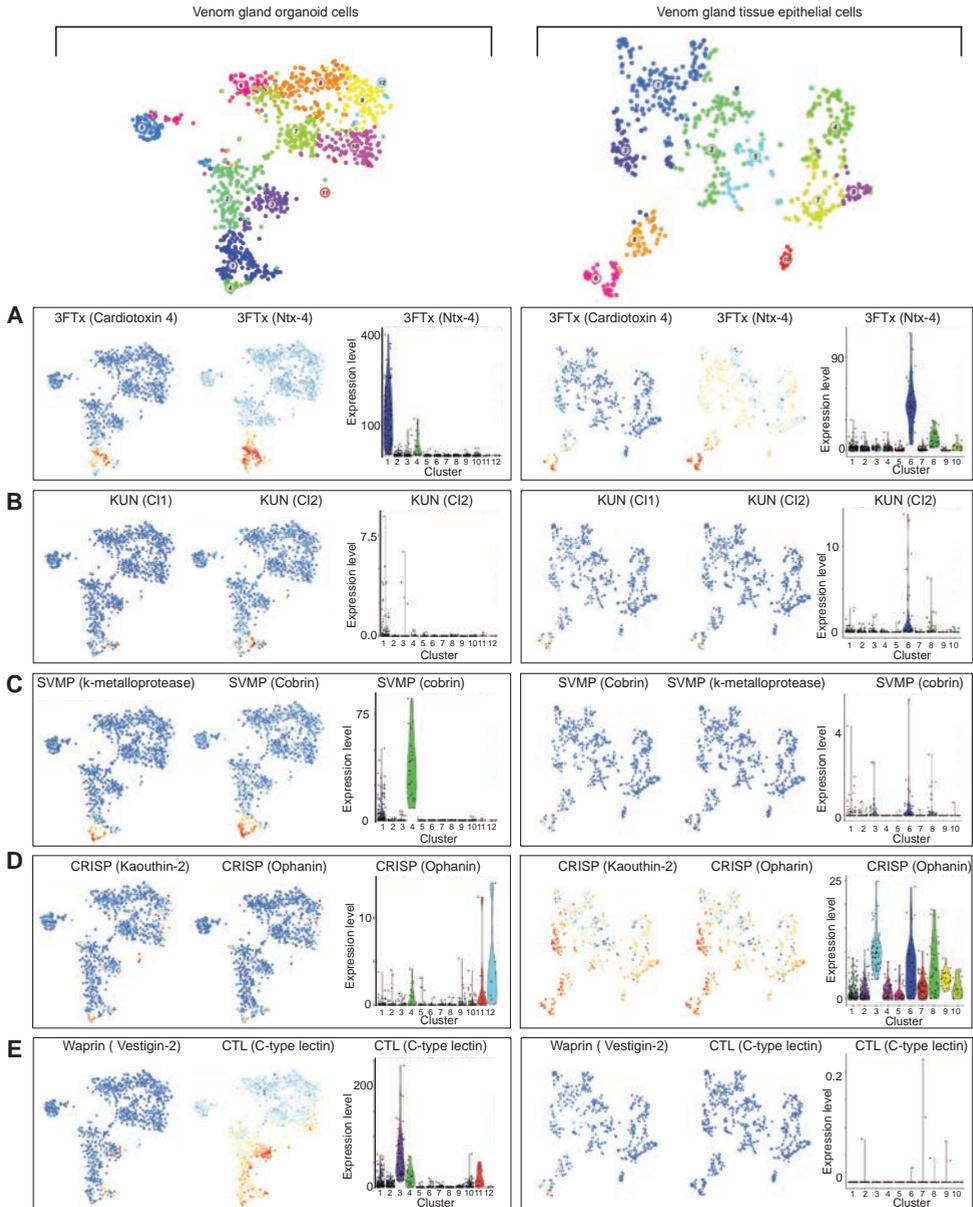
Figure S3. Dedicated venom gland epithelial cells express different classes of toxins, related to Figure 4.
(A-E) Expression levels of selected toxin genes in *t*-SNE map (color coded logarithmic scale of transcript expression) and violin plots visualizing expression levels of cluster-enriched toxins. Left panels are venom gland organoid cells and right panels are venom gland tissue epithelial cells, both from *A. l. cowlesi*.
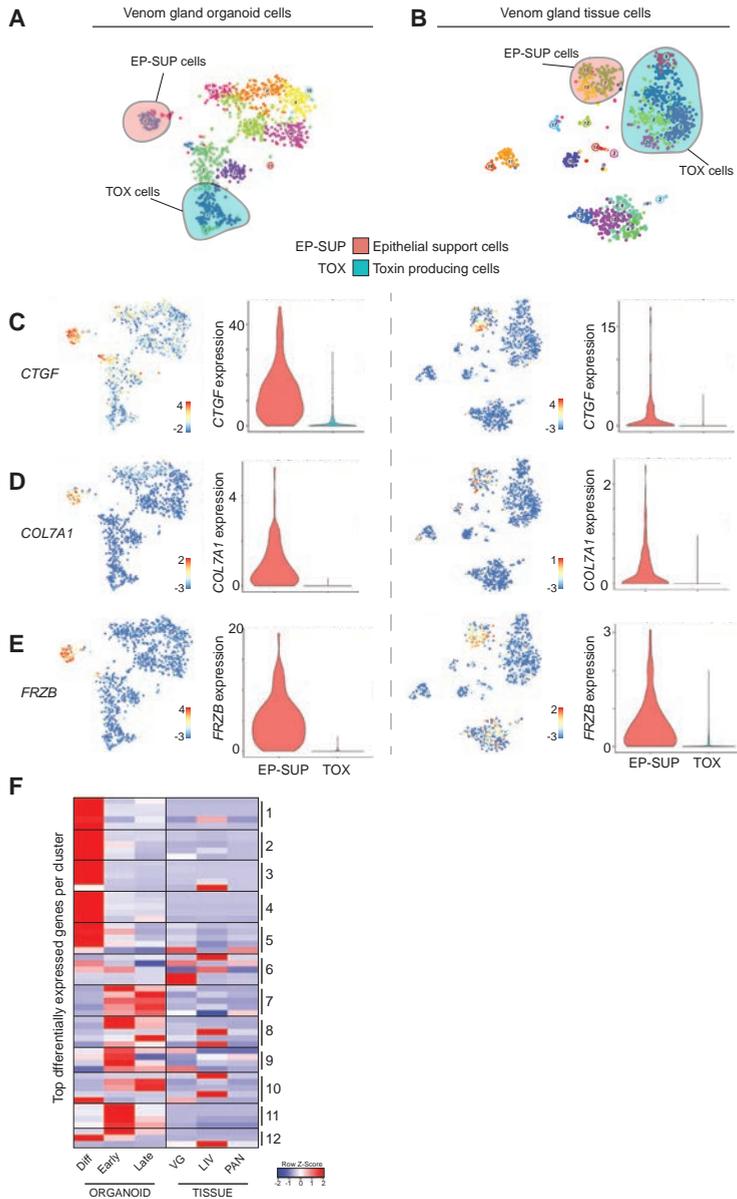
Figure S4. Two different lineages of epithelial cells captured in organoids and primary tissue, related to Figure 4. (A) *t*-SNE map of organoid dataset, highlighted are the clusters that account for the EP-SUP and TOX lineage. **(B)** *t*-SNE map of tissue dataset, highlighted are the clusters that account for the EP-SUP and TOX lineage. **(C-E)** Expression levels of selected EP-SUP genes in *t*-SNE map (color coded logarithmic scale of transcript expression) and violin plots visualizing expression levels in the two distinct lineages. Left panels are organoid cells and right panels are tissue cells. **(F)** Heatmap of marker expression (12 clusters (3 to 5 genes per cluster) from organoid single cell sequencing dataset (Figure 4)) in bulk mRNA samples from venom gland organoids in differentiation (Diff) and expansion medium early (Early) and late (Late) and venom gland (VG), liver (LIV) and pancreas (PAN) tissue. Expression levels visualized as color coded Row Z-score.

A Venom gland organoid cells
PLI

B Venom gland epithelial tissue cells
PLI

C *MKI67*

D *RNF43*

E Organoid   Tissue
ASCL2

F Organoid   Tissue
LGR5

G Venom gland organoid cells   Venom gland epithelial tissue cells
WNT10A

H *WNT9A*

I *FGF7*

J *CHRD*

K *FSTL1*

Epithelial support cells
Stromal cells
Toxin producing cells

L CTL *in situ* (slide 83)
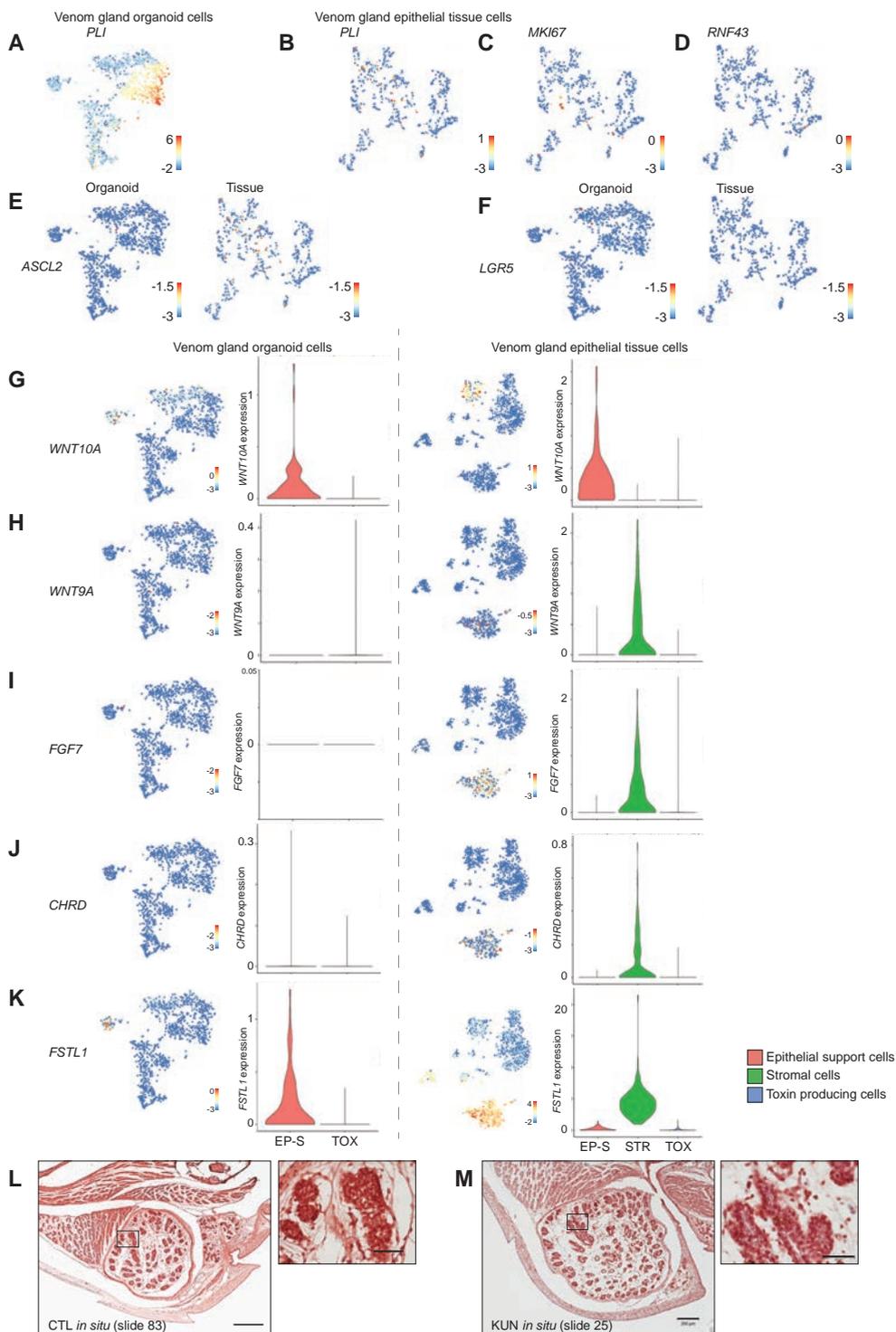
M KUN *in situ* (slide 25)

◀ **Figure S5. Non-venomous cells express stem cell markers and niche factors, regional heterogeneity in CTL and KUN expression *in vivo*, related to Figure 4, 5 and 6. (A-F)** Expression levels of selected genes in *t*-SNE map (color coded logarithmic scale of transcript expression). Left panels are organoid cells and right panels are tissue cells. **(G-K)** Expression levels of selected niche signal genes in *t*-SNE map (color coded logarithmic scale of transcript expression) and violin plots visualizing expression levels in the distinct lineages. Left panels are organoid cells and right panels are tissue cells. **(L-M)** Representative images of venom gland proximal and distal sections. BM-purple stain of *in situ* hybridization. Lack of detectable CTL (slide 83, distal) and KUN (slide 25, proximal) signal in *A. l. cowlesi* venom gland tissue. Scale bars, 200 μm, close up panels 50 μm.

A

NTX4_mNeonGreen fusion

NTX4 locus

Protein

Three-finger toxin

mNeonGreen
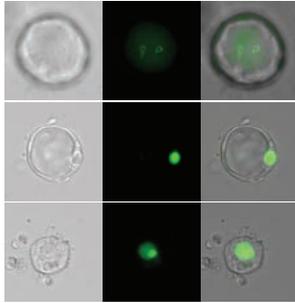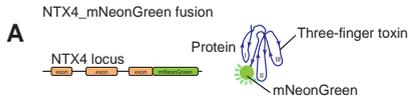
B

Hits mascot search engine:

3L21_ASPSC
Long neurotoxin 1 OS=Aspidelaps scutatus PE=1 SV=1
Protein score: 84
Sequence Coverage: 33%

3S13_NAJHA
Short neurotoxin 5 OS=Naja annulifera PE=1 SV=1
Protein score: 65
Sequence Coverage: 27%

CRVPA_NAJHA
Cysteine-rich venom protein annuliferin-a (fragment) OS=Naja annulifera PE=1 SV=1
Protein score: 48
Sequence Coverage: 33%

C

D

384 well plate

PhaseGuide™

PhaseGuide™

ECM gel

cells

E

Myotubes on Organoplate

F

Hoechst Desmin

Hoechst Dystrophin

G

Maximum projection calcium waves

alpha-Bungarotoxin

No carbachol

Vehicle control

I

Cell viability (%)

○ 24h
● 48h

ns

ns

Supernatant control organoids

Supernatant venom gland organoids

Alpha bungarotoxin

Control

H

Calcium wave propagation

alpha-Bungarotoxin (n=4)

No carbachol (n=3)

Vehicle control (n=9)

before   after

Carbachol stimulation

◀ **Figure S6. Venom gland organoids functionally secrete proteins that block calcium signals in muscle cells on OrganoPlate, Related to Figure 7. (A)** Example of protein hits obtained with a Mascot search of the spectra measured from LC-MS analysis. Based on tryptic digest of an organoid supernatant. **(B)** Generation of an endogenous fluorescent tag of three-finger toxin *NTX4* using CRISPR-HOT in *A. l. cowlesi* organoids **(C)** MIMETAS 2-lane OrganoPlate® **(D)** Schematic overview of OrganoPlate microfluidic network covering four wells of the 384-well-plate. **(E)** Representative immunofluorescent 3D reconstruction of myotubes forming a perfusable tube structure. Staining for Hoechst (blue) and Myosin (red). **(F)** Confocal image of mature myotubes stained for Desmin, Dystrophin (red) and Hoechst (blue). **(G)** Cultures were exposed to different supernatants in chips, after which calcium wave propagation was imaged before and after carbachol stimulation. Maximum projected images depict calcium wave propagation over all image frames after carbachol exposure, in which purple reflects activity to the left and green to the right. **(H)** Quantification of (F). The total calcium wave movement was calculated before and after carbachol stimulation based on total fluorescent signal. Data points represent biological replicates. **(I)** Cell viability assay (WST8) of C2C12 murine muscle cultures 24h and 48h after a 30-minute exposure to supernatants. Data points represent biological replicates.
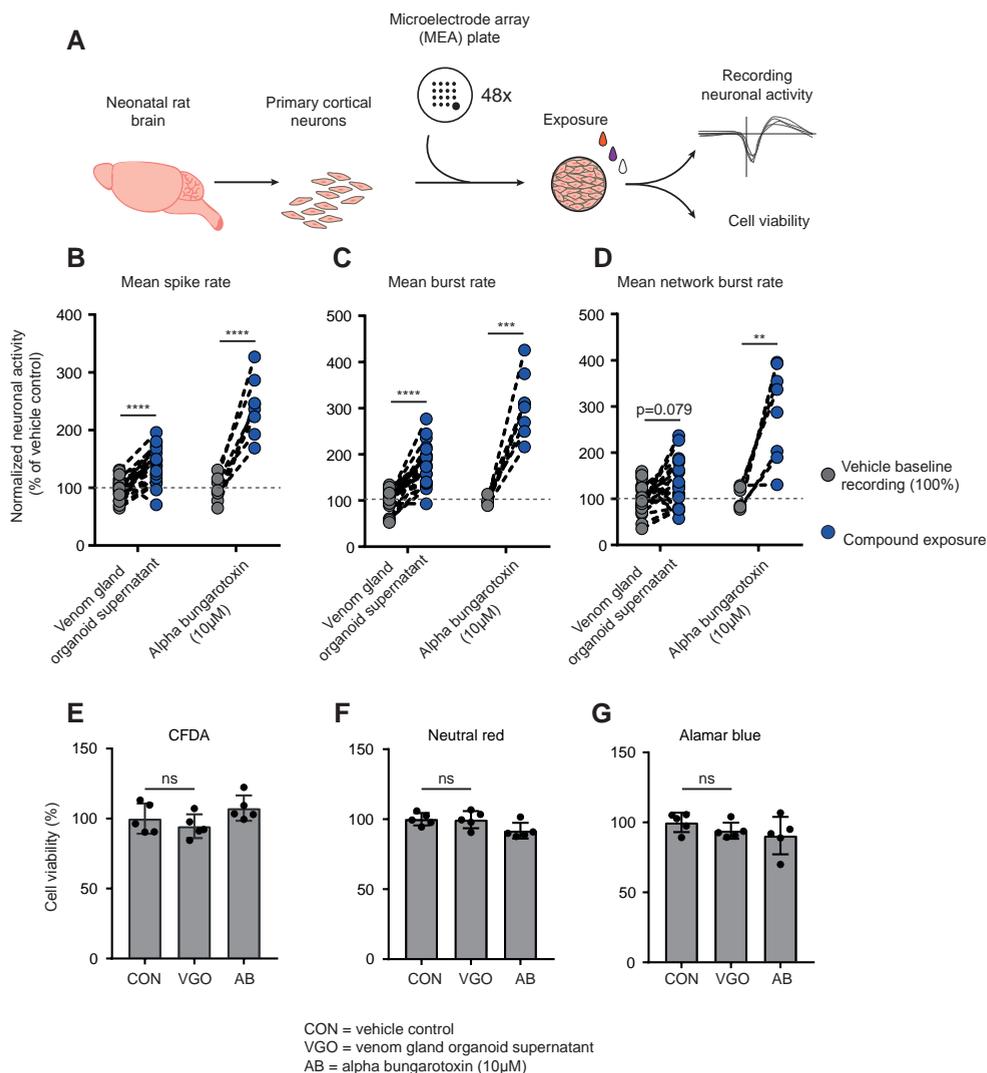
Figure S7. Venom gland organoid proteins increase neuronal activity in rat cortical neurons, related to Figure 7. (A) Schematic representation of experimental setup. Culture of neonatal rat cortical neurons on MEA plates. Exposure to organoid proteins and control is followed by recording of neuronal activity and cell viability assessment. **(B-D)** Quantification of neuronal activity measured as 'mean spike rate' (B), 'mean burst rate' (C) and 'mean network burst rate' (D) normalized to culture well-matched vehicle control. Exposure to venom gland organoid supernatant and recombinant alpha bungarotoxin. Paired t-test: **** p<0.0001, *** p=0.0002, ** p=0.005. Data points represent biological replicates. **(E-G)** Cell viability assessment using CFDA, Neutral red and Alamar blue. Normalized to vehicle control (100%). Differences are not significant (n.s.). Data points represent biological replicates.

SUMMARIZING DISCUSSION

5

Yorick Post[1]

[1] Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences (KNAW) and
UMC Utrecht, 3584 CT Utrecht, the Netherlands.

INTRODUCTION

Adult stem cells are cells in the body that have the capacity to replace lost tissue through cell division (Chapter 1) (Post and Clevers, 2019). The stem cell field was able to reach this one and only generalizable definition after several decades of research and development of better model systems. We now understand that adult stem cells exhibit widely divergent mechanisms and phenotypes to replace lost tissue in a given organ. In addition, one organ can be populated by multiple adult stem cell populations, each responding to a different stimulus or replacing a specific subset of cells. Adult stem cell versatility highlights the importance to investigate every organ or stem cell population as individual study object rather than apply a stereotypic template to all adult stem cell biology.

Cells and tissues are lost through natural physiological turnover by the wear and tear of the organ. Some cells, like the epithelial lining of the intestine, are replaced approximately every five days (Barker et al., 2007). Cells from such organs are often exposed to external insults and are able to dispose of their dead cells directly to the outside world. High turn-over organs require an active proliferative compartment to keep up with the constant loss of cells. Other organs such as the heart, especially the muscle cells; cardiomyocytes, have very low or neglectable turn-over rates in homeostatic conditions in the adult (Eschenhagen et al., 2017). Low turn-over organs consist of long-lived cells and the demand for a proliferative population to replace these cells is therefore lower. The endogenous regenerative capacity of adult stem cells declines during adulthood, resulting in phenotypes associated with natural ageing and age-related disease. Cells with the capacity to proliferate do carry the risk to propagate disease causing mutations across an organ, for example in the case of cancer. Oncogenic alterations in an adult stem cell are of higher risk in fueling a cancer than a transformed terminally differentiated cell that will be expelled from the system.

Upon tissue damage, the cellular population balance within an organ can dramatically change. Because of a changing chemical niche, or different mechanical signals, proliferation is increased to compensate for the loss of cells. The type of damage can be of influence on the cellular damage response, depending on the cells are directly affected and the severity of the insult. After the damage response, when cell numbers are restored, the proliferative activity returns back to normal. Of note, only a limited number of cell types will have the capacity to proliferate upon damage and take up the role as an adult stem cell. A cell's potential to proliferate upon damage, and repopulate the organ, can cause an irreversible shift in cellular composition after damage exposure. In addition, a proliferating adult stem cell will have a limited repertoire of cell types it will be able to differentiate into or even be unipotent. The differentiation potential will always be within the resident cell types of the organ of origin.

A better understanding of adult stem cells and their potential progeny is essential to interpret homeostatic organ function, cellular turnover and adult stem cell involvement in disease. After defining adult stem cell definition in *Chapter 1*, we applied different experimental models to tissues with remaining open questions. The existence of cardiac stem cells and the heart's response to damage in *Chapter 2*. Adult stem cell derived organoids and a single

cell atlas of the human lacrimal gland to model tear secretion and composition in *Chapter 3*. We tested the applicability of mammalian models to reptilian adult stem cells and characterize the snake venom gland in *Chapter 4*. Here, in *Chapter 5*, we will summarize our findings, discuss the implications these results and explore potential next steps for future research.

PROLIFERATIVE CELLS IN THE HEART

Over the past fifteen years, intensive research on the adult mammalian heart's capacity to regenerate itself has led to the general agreement that throughout life a small fraction of cardiomyocytes is replaced (Eschenhagen et al., 2017). The cellular turn-over rate and the cell source of new cardiomyocytes is still heavily debated. While some reports indicate the existence of an adult cardiac stem cell that is multipotent and support myocardial regeneration (Ellison et al., 2013), others deny such endogenous repair mechanism and describe an extremely low turnover of already existing cardiomyocytes (Li et al., 2018). Previous studies relied heavily on stem cell markers taken from unrelated tissues (SCA-1, c-KIT), or used 'surrogate' stem cell characteristics such as DNA label retention. To address this issue, we have not restricted ourselves by preformed concepts, but rather applied the broadest and most direct definition of stem cell function: the ability of any cell to replace lost tissue by cell division, described in chapter 1.

In chapter 2, we have generated a map of proliferating cardiac cells and their progeny in growing murine neonatal hearts as well as homeostatic and injured murine adult hearts using single-cell messenger RNA-sequencing and genetic lineage tracing (Kretzschmar et al., 2018). For our mouse models, we have exploited the gene that is by far the most widely used proliferation marker in basic cell and developmental biology, in clinical oncology and in pathology: Ki67. With these models (Ki67-RFP and Ki67-Cre-ERT2) we marked every cell in the heart that becomes proliferative and were able to comprehensively determine the offspring. With single-cell sequencing of the Ki67-RFP mouse we purified and tracked down the identity of all cells in the heart that were in G1, S, G2 or M stage of the cell-cycle. In combination with the Ki67 lineage tracing mouse Ki67-Cre-ERT2 crossed with Rosa26-LSL-tdTomato we enabled the tracing of all proliferating cells at the time of tamoxifen injection with a fluorescent marker. The benefit of these models is that we were not limited or biased by cell type specific markers. If any new cardiomyocyte was formed, through proliferation, we would detect it. The downside would be that if we were to find traced tdTomato-positive cardiomyocytes we would not directly know its cellular origin. In other words, to distinguish between a proliferating cardiomyocyte or a proliferating stem cell that would differentiate into a cardiomyocyte during the time of the tracing experiment.

We found that -while neonatal cardiomyocytes are actively cycling- no significant *de novo* generation of cardiomyocytes occurs from any cell source in the homeostatic or regenerating adult murine heart. Excluding the generation of new cardiomyocytes in the adult damaged heart 1) from a quiescent CSC population that re-enters the cell cycle, 2) by transdifferentiation of fibroblasts towards cardiomyocytes, or 3) for increased cell cycle entry of cardiomyocytes in

response to cardiac injury. Of note, we measure this outcome at the specific timepoints post injury and the type of damage models we used in this study. This is in line with other reports, describing the proliferation of cardiomyocytes to be limited to a number of days after birth (Li et al., 2018). In the neonatal hearts, pre-existing cardiomyocytes function as adult stem cells by re-entering the cell cycle upon damage (Porrello et al., 2011). A few animals in nature, such as the zebrafish, retain this cardiac regenerative capacity as an adult and is therefore a much-used model organism (Poss et al., 2002).

The loss of cardiomyocytes turned out to be compensated by a strong increase in fibroblast proliferation. Injury activated cardiac fibroblasts acquired a gene-expression pattern similar to that of neonatal cardiac fibroblasts to form a fibrotic scar. Using lineage tracing, we demonstrate that this cardiac fibroblast population is marked by follistatin-like 1 (Fstl1). Fstl1-positive fibroblasts function as unipotent adult stem cells in the damage response of the heart. Without autocrine Fstl1 signalling (a BMP antagonist) a scar is not formed and the heart ruptures. While a non-contractile scar does not seem beneficial to cardiac function, it is actually crucial in the survival of a cardiac infarct.

We believe that our data presented in this chapter serves as a comprehensive atlas for cardiac cell proliferation. This dataset can serve as an unbiased resolution in the cardiac stem cell debate. One could always argue more cells should have been sequenced or more experimental conditions tested. However, we believe if there was any significant contribution of *de-novo* cardiomyocytes in the adult heart, we would have detected it with the experimental setup presented here. The results we observe in mice should be validated in human tissue in the future. Soberingly, putative cardiac stem cells are still used in clinical trials today with patients suffering from myocardial infarction. More fundamental and clinical research should be devoted to the importance of fibrotic scar formation.

THE CELLULAR ORIGIN OF TEARS

Adult stem cell-derived organoids have been established of almost every epithelial organ. Examples include the intestinal lining (Sato et al., 2009), liver ductal cells (Huch et al., 2015) and hepatocytes (Hu et al., 2018), pancreas ductal cells (Boj et al., 2015), salivary gland (Maimets et al., 2016) and many more (Clevers, 2016). In these cultures, it is always the epithelial compartment that can be long-term expanded, surrounding tissue such as fibroblasts, endothelial and immune cells are lost during the passaging. The organoids are a robust model to study organ specific epithelial biology. Examples of such studies include the elucidation of differentiated cell lineages from an adult stem cell and their secreted products, either endocrine or exocrine (Beumer et al., 2018). One of the secretory epithelial tissues that is not completely understood and was lacking a robust organoid model is the lacrimal gland, also known as the tear gland.

In Chapter 3 of this thesis we set out to develop and characterize an organoid model of the human lacrimal gland. A Complete understanding of this organ in both healthy and diseased conditions is lacking, compared to other glandular organs, due to absence of insight into the cellular composition. It is largely unknown what adult stem cell populations are present

in the adult human lacrimal gland and what the function of their progeny is. As for almost all the other organoid cultures coming from our lab we started with mouse tissue. Murine lacrimal gland cells can be long-term expanded as organoids in a generic organoid growth factor medium of high Wnt activation, activation of FGF signaling and the inhibition of TGFbeta and BMP signaling. The organoids will self-organize as the classical cystic structures. The lacrimal gland origin and identity of these cultured cells was confirmed by nuclear expression of Pax6. The contamination of neighboring tissues during the dissection can be problematic in the expansion of organoid cultures, especially since culture media of different organs are rather similar. In the case of the lacrimal gland we had to rule out any outgrowth of salivary gland tissue given its anatomical location. Mouse lacrimal gland organoids expressed several tear components found *in vivo* and had the potential to differentiate upon the withdrawal of most growth factors. The withdrawal of growth factors from expansion medium mimics the absence of niche factors *in vivo* and results in the exit of the cell cycle and the start of differentiation. The timing and components added to the differentiation can be of influence on the cellular outcome (Beumer et al., 2018), this was not used in the murine lacrimal gland differentiation presented in Chapter 3. We next showed the applicability of CRISPR-Cas9 genome editing in organoids by knocking out *Pax6*, the master regulator of eye lineages (Makarenkova et al., 2000). While the loss of transcription factor Pax6 did not affect cell viability, it blocked cells from fully differentiating upon exposure to differentiation medium. The use of genome editing techniques such as CRISPR-Cas9 mediated frame shifts enables genotype-phenotype studies in adult epithelial organs. To do this for Pax6 *in vivo* would require a different approach because of its role during embryonic development (Makarenkova et al., 2000).

The success we had with mouse lacrimal gland tissue encouraged us to attempt human lacrimal gland tissue as well. Surplus material from human lacrimal gland biopsies was successfully expanded as organoid for multiple passages. The human organoids grew slightly slower and had a more differentiated phenotype in expansion conditions compared to mouse. This can be due to intrinsic properties of the human epithelium, non-optimal growth medium or the condition of the tissue when the biopsy was taken. While the latter option would be interesting to link disease status to organoid phenotype, at this point we were not able to do that as the biopsy material is anonymized and disease status confidential. Since the organoid cells formed a polarized epithelium with the apical side towards the lumen of the organoid we developed a function secretion assay. We used functional organoid swelling (by increased luminal volume) as a proxy for tear secretion. With this assay we identified noradrenaline and a combination of FSK/PGE2 as potent stimulators of secretion. Future research can exploit this assay to identify novel inducers of tear secretion or study patient-derived organoids to better understand the effect of disease genotype on tear secretion. However, dry eye disease and Sjogren's syndrome are disorders described to have an autoimmune background in the progression of the disease. In these cases, it will be relevant to explore organoid co-cultures with of resident immune cells, as previously illustrated with tumor-derived organoids.

We present the organoid dataset in this chapter together with a single-cell atlas of the human lacrimal gland. In all tissues it is very valuable to have a reference map of all the cells present

in vivo. Single-cell sequencing revealed the spread of individual tear components over ductal cells and acinar cells. Even within the ductal cells, heterogeneity exists in the most abundantly expressed tear components. This is the first transcriptional description of cell types for the human lacrimal gland. Of note, the lacrimal gland organoids, from both mouse and human, recapitulate the ductal compartment of the gland. This is also observed in organoids from pancreas and salivary gland tissue (Boj et al., 2015; Maimets et al., 2016). Future research should be focused on the expansion of acinar cells specifically, as it most likely requires a different protocol from the start of the cultures rather than the ability to differentiate ductal cells into acinar cells.

The research presented in this chapter provides an experimental platform to study the (patho-)physiology of the tear gland.

FROM STEM CELLS TO VENOM

In the last decade organoids have revolutionized our toolbox to investigate development and disease of almost every human organ in the body (Clevers, 2016). The adult stem cell-derived organoid story took off with the establishment of murine small intestinal organoids in 2009 (Sato et al., 2009). This was the first demonstration of long-term culture conditions that enable the formation of a 'mini-organ' *in vitro* from a single adult stem cell, marked by Lgr5. The mouse small intestinal organoids contain both stem cells and differentiated cells, and is up to date still one of the most complete organoid models published. The establishment of mouse organoids finds its origin in the easy accessibility to tissue and the conservation in organ biology between mouse and human.

The research presented in chapter 4 originated from the question if we can expand organoid technology beyond mouse and human. There are at least two arguments to do this, 1) to study unique features of an organ that is shared between a model organism of interest and humans. For example, develop an organoid model of the regenerating adult zebrafish heart, to gain insight for a human clinical intervention. And 2) to use organoid technology to study an organ of interest that is not present in mouse or human. We decided to go with argument two, the snake venom gland. Venomous snakes are a major health threat to humans and kill more than 100,000 people every year. Yet, snakes are also an incompletely explored gold mine for derivation of pharmacological drugs. One of the first anti-hypertension drugs for example was isolated from venom of the Brazilian pit viper. Many snakes are difficult to access and study due to their rarity or large variation in toxin composition between individuals. This poses challenges on harnessing the full potential of their pharmacological gold mine, but also for the production of anti-venom.

We started the establishment of venom gland organoids off with late-embryonic tissue. We applied a nearly identical protocol our lab has used for different mouse and human organs; the digestion of tissue in small cellular fragments and plate these in a growth factor rich medium embedded in a matrix. The first hurdle to overcome was if the mammalian protein-based culture medium was able to activate the cells from the reptilian venom gland. This is crucial in organoids cultures as you want to activate the resident cells to proliferate rather than

just cultivate surviving primary tissue. It turned out the reptilian receptor - mammalian ligand interaction is conserved enough for the components in our medium (FGF10, EGF, R-spondin, Noggin, etc.) to activate the snake cells.

Using this protocol, we established long-term expanding venom gland organoids from several snake species. We focused our characterization on organoids from *Aspidelaps lubricus*, the Cape coral snake. Since there was no Cape coral snake genome published we generated a newly assembled *de-novo* transcriptome for the characterization of our organoids. This is something to keep in mind shifting to different animals for organoid cultures. Of course, one could pick a more characterized organism with more resources as a model of interest. The Cape coral snake organoids revealed to express high levels of toxin transcripts. As in the previous chapters, we set out to use single-cell mRNA sequencing to reveal the cellular composition of the venom gland and the organoids. The sequences were mapped on our new transcriptome. We identified distinct venom-expressing cell types as well as proliferative cells expressing homologs of known mammalian stem cell markers. Also, a hard-wired regional heterogeneity in the expression of individual venom components is maintained in organoid cultures. Venom peptides are complex in their tertiary structure and require several enzymes for correct folding.

To determine whether the organoid cultures produced functionally active venom components we developed a protocol to harvest the luminal content of the organoids. Using LC-MS we determined that the harvested peptides reflect crude venom composition. Exposing murine muscle cells and rat cortical neurons to these peptides revealed the 'organoid venom' to be functionally active by disruption calcium signal propagation and control of spontaneous action potential firing. The production and secretion of functionally active venom components opens up several avenues for follow up research. For example, the production of venom in a chemically defined environment is of high interest in antivenom production, which is at the moment hampered by incomplete understanding of intraspecies heterogeneity of venom composition. Other *in vitro* systems used for protein expression and purification are not able to translate functionally active venom peptides upon the introduction of the gene, due to the complex machinery involved. Another research angle would be to study the mechanism by which venom gland cells control exposure to their own venom, as an endogenous antivenom. We have not noticed any effect of autocrine toxicity in our venom gland organoids. Some cells in the single-cell sequencing dataset transcribe genes that are described to have neutralizing properties against some venom components. This would require further validation. Of note, most successful and robust organoid propagation and characterization was performed with Elapid species. Viperidae snakes are described to have a more cytotoxic venom composition, potentially harming itself if not kept under control. It will be worthwhile to further investigate whether viper derived organoids produce functionally active cytotoxic components and how these are controlled. The increased applicability of genome editing tools such as CRISPR-Cas9 also reached the field of organoids in recent years. We demonstrated the use of CRISPR-HOT in tagging a three-finger toxin in the Cape coral snake organoids. With the potential to apply changes to the genome or introduce exogenous genes, opportunities are endless. Of interest can be the expression of several medically relevant toxins (from different species) in one organoids

line. This "pan-venom" can be used for antivenom production. Or the introduction of point mutations in toxins to alter or specify its human target receptor or cell in an approach to pharmaceutically design potent molecules.

In short, the research presented in Chapter 4 extends organoid technology to reptilian tissues and describes an experimentally tractable model system representing the snake venom gland. It opens up avenues to exploit these cultures to better understand the potential and danger of venom and to apply organoid technology to other animal species of interest

1

2

3

4

5

REFERENCES

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., et al. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449, 1003-1007.

Beumer, J., Artegiani, B., Post, Y., Reimann, F., Gribble, F., Nguyen, T.N., Zeng, H., Van den Born, M., Van Es, J.H., and Clevers, H. (2018). Enteroendocrine cells switch hormone expression along the crypt-to-villus BMP signalling gradient. Nat Cell Biol 20, 909-916.

Boj, S.F., Hwang, C.I., Baker, L.A., Chio, II, Engle, D.D., Corbo, V., Jager, M., Ponz-Sarvise, M., Tiriac, H., Spector, M.S., et al. (2015). Organoid models of human and mouse ductal pancreatic cancer. Cell 160, 324-338.

Clevers, H. (2016). Modeling Development and Disease with Organoids. Cell 165, 1586-1597.

Ellison, G.M., Vicinanza, C., Smith, A.J., Aquila, I., Leone, A., Waring, C.D., Henning, B.J., Stirparo, G.G., Papait, R., Scarfo, M., et al. (2013). Adult c-kit(pos) cardiac stem cells are necessary and sufficient for functional cardiac regeneration and repair. Cell 154, 827-842.

Eschenhagen, T., Bolli, R., Braun, T., Field, L.J., Fleischmann, B.K., Frisen, J., Giacca, M., Hare, J.M., Houser, S., Lee, R.T., et al. (2017). Cardiomyocyte Regeneration: A Consensus Statement. Circulation 136, 680-686.

Hu, H., Gehart, H., Artegiani, B., C, L.O.-I., Dekkers, F., Basak, O., van Es, J., Chuva de Sousa Lopes, S.M., Begthel, H., Korving, J., et al. (2018). Long-Term Expansion of Functional Mouse and Human Hepatocytes as 3D Organoids. Cell 175, 1591-1606 e1519.

Huch, M., Gehart, H., van Boxtel, R., Hamer, K., Blokzijl, F., Verstegen, M.M., Ellis, E., van Wenum, M., Fuchs, S.A., de Ligt, J., et al. (2015). Long-term culture of genome-stable bipotent stem cells from adult human liver. Cell 160, 299-312.

Kretzschmar, K., Post, Y., Bannier-Helaouet, M., Mattiotti, A., Drost, J., Basak, O., Li, V.S.W., van den Born, M., Gunst, Q.D., Versteeg, D., et al. (2018). Profiling proliferative cells and their progeny in damaged murine hearts. Proc Natl Acad Sci U S A 115, E12245-E12254.

Li, Y., He, L., Huang, X., Bhaloo, S.I., Zhao, H., Zhang, S., Pu, W., Tian, X., Li, Y., Liu, Q., et al. (2018). Genetic Lineage Tracing of Nonmyocyte Population by Dual Recombinases. Circulation 138, 793-805.

Maimets, M., Rocchi, C., Bron, R., Pringle, S., Kuipers, J., Giepmans, B.N., Vries, R.G., Clevers, H., de Haan, G., van Os, R., et al. (2016). Long-Term In Vitro Expansion of Salivary Gland Stem Cells Driven by Wnt Signals. Stem Cell Reports 6, 150-162.

Makarenkova, H.P., Ito, M., Govindarajan, V., Faber, S.C., Sun, L., McMahon, G., Overbeek, P.A., and Lang, R.A. (2000). FGF10 is an inducer and Pax6 a competence factor for lacrimal gland development. Development 127, 2563-2572.

Porrello, E.R., Mahmoud, A.I., Simpson, E., Hill, J.A., Richardson, J.A., Olson, E.N., and Sadek, H.A. (2011). Transient regenerative potential of the neonatal mouse heart. Science 331, 1078-1080.

Poss, K.D., Wilson, L.G., and Keating, M.T. (2002). Heart regeneration in zebrafish. Science 298, 2188-2190.

Post, Y., and Clevers, H. (2019). Defining Adult Stem Cell Function at Its Simplest: The Ability to Replace Lost Cells through Mitosis. Cell Stem Cell 25, 174-183.

Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., et al. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 459, 262-265.

NEDERLANDSE SAMENVATTING

CURRICULUM VITAE AND
LIST OF PUBLICATIONS

DANKWOORD / ACKNOWLEDGEMENTS

&

NEDERLANDSE SAMENVATTING
(ook voor niet-ingewijden)

MODELLEREN VAN GEWERVELDEN VOLWASSEN STAMCELLEN EN HUN NAKOMELINGEN
Yorick Post

Het onderzoek beschreven in dit proefschrift is gericht op de ontwikkeling en het gebruik van modelsystemen voor volwassen stamcellen. Ons lichaam en alle organen zijn opgebouwd uit cellen, deze bouwstenen zijn ongeveer een honderdste van een millimeter groot en we hebben er zo'n 40 biljoen. Volwassen stamcellen bevinden zich in bijna ieder orgaan van een volwassen mens. Deze cellen zijn verantwoordelijk voor het onderhoud en herstel van onze organen. Dat doen ze door zichzelf te delen, van één naar twee cellen, de nakomelingen. Met deze verdubbeling kunnen verloren of beschadigde cellen in ons lichaam worden vervangen. Een volwassen stamcel beperkt zich tot de grenzen van het orgaan waar het gevonden kan worden; een darmstamcel maakt alleen maar darmcellen, een leverstamcel maakt alleen maar levercellen, enzovoort. Het natuurlijk verlies van cellen en weefsel gebeurd in ieder orgaan, echter wel op verschillende snelheden. Een extreem voorbeeld is de darm, alle cellen die de binnenkant van onze darm bekleden worden zo'n iedere vijf dagen compleet vervangen. De stamcellen in de darm vermenigvuldigen zich actief om voor de constante toevoer van nieuwe cellen te zorgen. Aan het andere uiteinde van het spectrum vind je de spiercellen van het hart. De spiercellen waarmee je geboren wordt doe je bijna de rest van je leven. Daarom kan een hartinfarct, schade aan het hart, levensgevaarlijk zijn en langdurige gevolgen hebben. De aanwezigheid van stamcellen of cellen die zichzelf de rol van een stamcel kunnen opleggen bepaalt vaak ook hoe snel en efficiënt een orgaan zichzelf kan repareren na schade. Denk bijvoorbeeld aan de stamcellen in je huid als je jezelf snijdt of de stamcellen in de lever als je een avond te veel alcohol hebt gedronken. Alhoewel ieder orgaan en ieder dier dit op een verschillende snelheid en met een andere strategie doet, het verlies en herstel van cellen vindt plaats in het hele dierenrijk.

Een beter begrip van volwassen stamcellen is cruciaal voor het begrijpen hoe organen zich in stand houden in gezondheid maar ook tijdens ziekte. In de laatste tientallen jaren heeft het volwassen stamcel onderzoek een vlucht genomen. De ontwikkeling van steeds betere onderzoeksmethoden heeft ons de gelegenheid gegeven om vrijwel alle cellen van ieder orgaan in kaart te brengen. Dus zowel de stamcellen als hun nakomelingen. In het onderzoek van dit proefschrift maken we gebruik van twee verschillende technieken om stamcellen te bestuderen in verschillende dieren en organen.

De eerste is *Lineage Tracing*, letterlijk vertaald 'afkomst traceren'. Met deze techniek kunnen we een cel in een muis een markering of een kleur geven. Deze markering is vaak een chemische stof die we kunnen kleuren of een fluorescerend eiwit dat oplicht onder de microscoop. Als deze cel zichzelf deelt en vermenigvuldigt, zoals een stamcel doet, hebben de nakomelingen van deze cel ook deze markering. Gebaseerd op de DNA-sequentie, de genetische code, kun je

heel nauwkeurig bepalen met welke cel je het traceren wilt beginnen. Op deze manier kunnen we stamcellen identificeren en precies volgen wat voor nakomelingen ze maken als ze zichzelf delen. Dit kunnen we doen in gezonde omstandigheden tijdens het ouder worden van de muis, of onderzoeken hoe stamcellen reageren op schade aan het orgaan. In dit proefschrift traceren we cellen met een fluorescerend eiwit in gezonde muizen en muizen met een hartinfarct.

De tweede techniek is *Organoid Technology*, ofwel het kweken van organoïden. Organoïden worden ook wel mini-organen genoemd. Het zijn orgaan kweekjes die ontstaan uit volwassen stamcellen en hun nakomelingen. Als we de volwassen stamcellen in kweek brengen onder de juiste omstandigheden, vormen ze driedimensionale structuren bestaande uit cellen die zichzelf organiseren. Dit klompje van cellen, de organoïd, bevat de cellen van het orgaan en bereikt een grootte van ongeveer een halve millimeter. Vandaar de naam mini-orgaan. Deze techniek werd als eerste toegepast op de darm, darm organoïden, inmiddels zijn er lever, alvleesklier, maag, slokdarm, mondholte, nier, long, etc. organoïden. Net zoals "biologische wetten en regels" in het lichaam, zal een darmstamcel alleen maar een darm organoïd maken, een leverstamcel enkel een lever organoïd, etc. Ondanks dat het hele dierenrijk volwassen stamcellen bevat, zijn vrijwel alle beschreven organoïden verkregen van de muis en mens. Omdat de organoïden bestaan uit volwassen stamcellen en het orgaan zowel structureel als functioneel benaderen vormen ze een sterk en makkelijk hanteerbaar model. Het onderzoek in dit proefschrift maakt gebruik van volwassen stamcel verkregen organoïden van verschillende organen en dieren.

In de volgende hoofdstukken gaan we opzoek naar stamcel identiteit, stamcel functie en de rol van stamcel nakomelingen in verschillende weefsels van de muis, mens en slang. Gebruikmakend van lineage tracing en organoid technology.

In **Hoofdstuk 1** formuleren we de meest generaliseerbare definitie van een volwassen stamcel: het vermogen om verloren weefsel te vervangen door celdeling. De eerste wetenschappelijke onderzoeken die stamcellen beschreven waren vooral gebaseerd op de principes van de hematopoëtische stamcel van het bloed. We beschrijven de meest recente onderzoeken en conclusies in verschillende organen en vergelijken het met het oorspronkelijke model. Al het bewijs wijst erop dat stamcellen bijzonder heterogeen zijn en in velen vormen en maten komen. Elk weefsel lijkt zijn eigen cellulaire strategie te hebben. Toegewijde stamcellen kunnen zeldzaam zijn, in veelvoud voorkomen als voorgangers van het uiteindelijke celtype of zelfs als gedifferentieerde cel met de potentie weer een stamcel te worden. Daarom kan een stamcel enkel gedefinieerd worden aan de hand van zijn functie om weefsel te herstellen door celdeling.

In **Hoofdstuk 2** passen we de definitie van een stamcel, beschreven in Hoofdstuk 1, toe op de vraag of het hart beschikt over een volwassen stamcel. Het bestaan van hartstamcellen in het hart heeft altijd onder zware discussie gestaan en het onderzoek gebaseerd op karakteristieken van andere weefsels. In dit onderzoek brengen we alle delende cellen in het neonatale, volwassen en beschadigde hart onpartijdig in kaart. Dit doen we aan de hand van een gen dat enkel tot expressie komt in delende cellen. Met het analyseren van al deze cellen (delend en niet-delend) komen we tot de conclusie dat er geen stamcel mechanisme bestaat voor de generatie van

nieuwe spiercellen. Echter, het littekenweefsel dat vormt na een hartinfarct is opgebouwd uit velen delende cellen, een reactie die cruciaal is om het hartweefsel bij elkaar te houden.

Hoofdstuk 3 beschrijft een nieuw organoïd model van de traanklier. Traanklier organoïden van de muis en mens kunnen langdurig worden vermenigvuldigd en lijken functioneel op het oorspronkelijke weefsel. De organoïden bestaan uit delende stamcellen en gedifferentieerde cellen die traancomponenten uitscheiden. De uitgescheiden eiwitten hebben veelal een antibacteriële functie, in het lichaam voorkomen deze componenten een ooginfectie. Al deze componenten hopen zich op aan de binnenkant van de organoïd en doen het zwellen als een ballon gevuld met tranen. In combinatie met een cel atlas van traanklier weefsel beschrijven we met de organoïden een nieuw experimenteel platform om de gezonde traanklier en mogelijke aandoeningen te bestuderen.

In **Hoofdstuk 4** breiden we organoïd technologie uit van muis en mens naar reptielen, de gifklier van een slang. Volwassen stamcellen uit de gifklier kunnen gegroeid worden als organoïden met het gebruik van menselijke groeifactoren. Deze factoren bootsen de signalen na die de cellen in het lichaam van de slang normaal ontvangen. De gifklier organoïden bestaan uit zowel delende stamcellen als verschillende gif-producerende cellen. Het gif kan uit de organoïden geoogst worden en is functioneel. We beschrijven dat verschillende cellen en regio's in de gifklier verschillende toxines produceren, dit fenomeen blijft bewaard en kunnen we nabootsen met de organoïden. De mini-gifklieren vormen een basis voor de productie van gif in het lab en opent de deur naar mogelijke toepassingen in de productie van antigif en gif-geïnspireerde medicijnen.

Hoofdstuk 5 vat de bevindingen van dit proefschrift samen en onderzoekt de implicaties voor toekomstig onderzoek naar volwassen stamcellen.

CURRICULUM VITAE

Yorick Post was born on the 17th of September 1992, in Den Haag, The Netherlands. In 2010, he graduated from Huygens Lyceum in Voorburg, majoring in Nature and Health and Nature and Technology tracks. He obtained a Bachelor of Science degree in Biomedical Sciences from the University of Amsterdam in 2013. During his bachelor's degree he conducted research at the Heart Failure Research Center at the Amsterdam UMC. In 2013, he enrolled in the research Master's program Biomedical Sciences, experimental oncology track, at the University of Amsterdam. His first internship was performed at the Netherlands Cancer Institute working on bladder cancer metastasis in the laboratory of Prof. Daniel Peeper. He spent his final internship at the Whitehead Institute, part of MIT, in Cambridge, United States. This research was supported by a Dutch Cancer Society Scholarship enabling him to work on genome-wide CRISPR screens to study human cancers in the laboratory of Prof. David M Sabatini. He graduated from the Master's program (*cum laude*) and obtained his Master of Science degree in 2015. That same year he started his PhD in the laboratory of Prof. Hans Clevers at the Hubrecht Institute in Utrecht, The Netherlands. The results of that research are printed in this thesis.

PUBLICATIONS RELATED TO THIS THESIS

Post, Y., Bannier-Helaouet, M., Bustos, M.T., Bar-Ephraim, Y.E., Van der Vaart, J., Begthel, H., Korving, J., Kalmann, R., Imhoff, S.M., and Clevers, H. (2020). A human tear gland atlas and organoids that can cry. *In progress*.

Post, Y., and Clevers, H. (2019). Defining Adult Stem Cell Function at Its Simplest: The Ability to Replace Lost Cells through Mitosis. *Cell Stem Cell 25*, 174-183.

Post, Y*., Puschhof, J*., Beumer, J*., Kerkkamp, H.M., de Bakker, M.A.G., Slagboom, J., de Barbanson, B., Wevers, N.R., Spijkers, X.M., Olivier, T*., et al.* (2020). Snake Venom Gland Organoids. *Cell 180*, 233-247 e221.
(Co-first authors)

Kretzschmar, K*., **Post, Y***., Bannier-Helaouet, M., Mattiotti, A., Drost, J., Basak, O., Li, V.S.W., van den Born, M., Gunst, Q.D., Versteeg, D*., et al.* (2018). Profiling proliferating cells and their progeny in damaged murine hearts. *Proc Natl Acad Sci* U S A *115*, E12245-E12254.
(Co-first authors)

Beumer, J., Artegiani, B., **Post, Y**., Reimann, F., Gribble, F., Nguyen, T.N., Zeng, H., Van den Born, M., Van Es, J.H., and Clevers, H. (2018). Enteroendocrine cells switch hormone expression along the crypt-to-villus BMP signalling gradient. *Nat Cell Biol 20*, 909-916.

DANKWOORD / ACKNOWLEDGEMENTS

Performing scientific research at the highest level is teamwork. Not only with the people that are directly involved with the experiments but also with the people that support you from the outside. The last four years that shaped this thesis and changed me as a person would not have been possible without this network of support. There are a number of people I wish to thank especially.

First of all, my promoter **Hans**. Thank you for giving me the opportunity to join your lab and sharing your vision on science with me. You have created an environment that brings highly motivated and incredibly skilled people with the same passion together. The output of this scientific pressure cooker speaks for itself. Early on in my PhD you told me to take more risk and explore areas off the beaten path. I think this piece of advice really paid off in the projects I pursued after and made me a more creative scientist. I admire your efficient way of passing on feedback. A one-minute conversation at the coffee machine sometimes determined the experiments I was going to do the weeks after. Because of the freedom and responsibility you give to all of your PhD students, we leave the lab with much more than just a publication list. I also really appreciate your help in launching my post-PhD career.

My paranymphs, friends and scientific brothers, **Joep** and **Jens**. The JYJ combination was definitely one of the highlights of my PhD. Others recognized so much unity in our work together they assumed my name was also written with a J: the three J's. Both of your hard work, drive to make big discoveries and intelligence were incredibly motivating. We really pushed each other to become better as individuals and as a team. But maybe even more important the fun we had together; the countless number of memes, pranks, unacceptable jokes, football in the hallway and office chair racing (I will not disclose which one) balanced everything out. This is what made the late evenings and weekends of work totally worth it and make the four years fly by. Bringing the snake venom gland organoids paper to a successful ending together and getting so much recognition for that work was the cherry on the cake! I am sure both of you will have a great future. Even though our paths are somewhat separated now I am confident the JYJ will be back in any shape or form!

Mijn lieve ouders, **Wouter** en **Kitty**. Zonder jullie onvoorwaardelijke steun over de jaren was dit boekje er niet geweest. Jullie vertrouwen in mij en de mogelijkheden die jullie mij gegeven hebben om in Amsterdam te studeren en een jaar in Boston door te brengen staan aan de basis van mijn carrière. Ik heb me nooit ergens zorgen om hoeven maken en heb me altijd volledig kunnen focussen op de dingen die ik graag deed. Met zo'n fijne thuisbasis om tot rust te komen en te herstellen kon ik altijd het maximale uit mezelf halen op werk.

My supporter through long- and short-distance, my beaner, **Lina**. Being on the other side of the Atlantic or next to me in Utrecht, you were always there for me. Thank you for always

checking on me if I was doing okay and motivating me. The prospect of living together at the end of my PhD definitely made me work harder and more efficient to make sure I didn't take too long. Spending my much-needed time off with you traveling together was amazing. Like a lucky charm you are always around when things work out: when I interviewed at the Clevers lab five years ago, when we put the first venom gland cells in culture and now we're going back together to defend this thesis.

All former and current Clevers PhD students I got to interact with. **Else,** jij was een fantastisch voorbeeld tijdens mijn PhD. Een ongelofelijk harder werker onder welke omstandigheden dan ook, en ook nog is zo druk buiten het lab. Bedankt voor adviezen, koffies en lol! **Maarten**, Marty, CRISPR-Cas-Marty, Als ik een pauze nodig had ging ik altijd even bij jou buurten. De Douwe Egberts machine heeft goed van ons geprofiteerd. Voorzichtig met je nerf gun. **Cayetano**, Caye it was great to see you grow from master student to legendary PhD student. Thanks for always throwing things at me, kept me sharp. Noh?**, Frans,** het was fijn om jou als senior PhD student te volgen toen ik begon. En als onderdeel van de old-days-coffee club. **Fjodor,** bedankt voor alle medische weetjes, culinaire tips en kidney inzichten. **Marie** aka Marie Blanpain, it was great to share the office with you. Your uncontrollable laughter made many of my days. Very happy we got to collaborate on the heart and the lacrimal gland projects! **Kim**, Zoetermeer – Nootdorp, wij snappen elkaar! Indrukwekkend hoe jij al je experimenten, collega's begeleiden, lab taken en een druk sociaal leven allemaal weet te combineren. **Jelte**, bedankt voor je immer kalmerende 'komt wel goed' mentaliteit. En voor het organiseren van alle PhD-outings! **Lulu**, you are an inspiration, moving to the other side of the world and jumping right into the new science and culture! **Kadi**, with Baltic discipline you kept all of the crazy PhD students (and Oded) on the right track. Thanks for that! **Carla**, thank you for keeping the much-needed balance in the other-side office. I am looking forward to your discoveries in the kidney field! **Evelyn** and **Margit**, our Princess Maxima Center additions! Thank you for your input during the lab meetings and representing the Clevers lab across the road.

Marc, the lacrimal gland project was more than just a master internship, you really put all your time, energy and passion into it. It was a pleasure to do this together and I'm looking forward to your future achievements.

All other former and current lab members**, Kai**, thank you for jump starting my PhD by involving me in the heart project. Thank you for sharing your vision on science and Germany with me. All the best with the Kretzschmar Lab! **Oded** aka Odko, we had so much fun! Sorry for calling your organoid cultures a graveyard. **Jarno**, it was an honor to learn culturing organoids from you! **Jasper** aka Jos, heerlijk om met jou het kantoor te delen. Ik weet niet of we er productiever van werden maar lol hadden we wel. **Janny, Wim, Karien, Stieneke, DJ, Sigrid, Benedetta, Helmuth, Nobu, Norman, Inha, Johan, Laura, Yotam, Jochem, Lena, Huili, Priyanca, Veerle, Anjali, Maaike, Harry, Jeroen, Carola, Yoshi, Tomo, Dymph, Talya, Amanda, Delilah, Georg,**