Hamilton Institute

Maynooth University
National University of Ireland Maynooth

Statistical analysis of high-dimensional spatio-temporal data

Student name: Chang Liu

Supervisor name: Prof. Ken Duffy

A thesis submitted in fulfilment of the requirements for the Ph.D. degree

at the

Hamilton Institute
Maynooth University
Maynooth, Co. Kildare, Ireland

February 18, 2025

Summary

To realise biological function, cellular behaviour is a dynamic, involved processes that encompass spatial arrangement, differentiation, division and death. Numerous *in vivo* biological experiments have been designed to track cellular behaviour over different time- and space-scale frames, generating a large amount of spatio-temporal data. However, traditional statistical methods are not well adapted to draw meaningful insight from those high-dimensional data. Motivated by primary data sources provided by collaborators, this thesis presents a novel statistical analysis framework of high-dimensional spatio-temporal data, addressing it at different time scales. To capture biological information at fixed time points, a spatial statistical pipeline is developed to quantify the distribution of various cell types and assess their spatial relationships. The analysis is extended to cover periods of the order of one day, with a statistical framework designed to process spatio-temporal data, combining a data cleaning process and investigating the relationship between cell movement and differentiation. Over longer time periods, mathematical models and statistical methods are developed to estimate the average number of divisions *in vivo*, offering insights into long-term cell distribution. The advanced statistical analysis helps capture the spatio-temporal relationships between different cell types, revealing the dynamic processes of cellular behaviour.

Acknowledgements

I would like to thank Prof. Ken Duffy for his invaluable time, understanding, patience and support throughout his supervision. His supervision has made me a better researcher and a better person.

I extend my thanks to my collaborators at Imperial College London and the Curie Institute: Sara, George, Flo, Cera and Lucie. Working alongside you has been a pleasure, and I truly appreciate your patience in helping me understand the complexities of biology. I am especially grateful to Prof. Cristina Lo Celso and Dr. Leïla Perié, whose scientific rigour has greatly advanced my research. You made me understand that teamwork makes the dream work.

I would like to thank my office mates for their constant moral support, you make my Irish life gorgeous. A special mention to Prof. David Malone, Prof. Andrew Parnell, Prof. Chris Brunsdon, Prof. Muriel Médard, Joanna, Rosemary, Kate and Janet for their tremendous encouragement and assistance throughout my PhD journey.

Finally, I would like to thank my mother and my father for their unwavering support.

Contents

List of Figures

List of Tables

1 Introduction

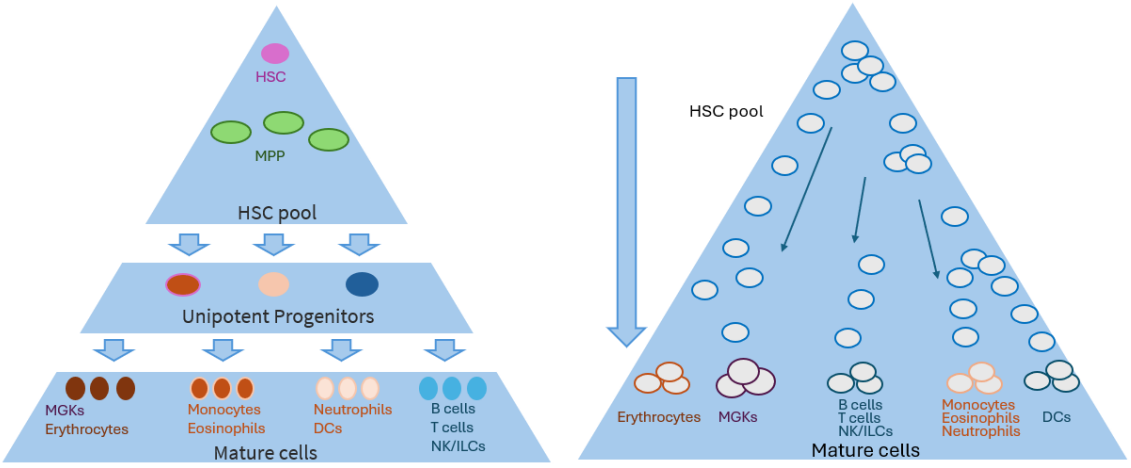
1.1 Abstract

This chapter introduces the essential concepts of haematopoiesis that are necessary to motivate the statistical framework developed in this thesis. It begins by introducing the haematopoietic system and its primary environment, the bone marrow; a tissue so packed with cells and it is often considered as a semi-solid tissue. The focus of this chapter is on summarizing the experimental methodologies that generate the types of data we will analyse later, particularly data on the spatial and temporal distribution of haematopoietic cells. Methods suitable for drawing inferences from complex data are central to this thesis. We describe current techniques for extracting and analysing data on cellular behaviour, which result from high-dimensional and complex datasets. The review highlights the challenges associated with studying dynamic processes such as cell movement, differentiation, and division in the context of haematopoiesis, emphasising the need for robust statistical frameworks and methods to handle the complexity of the data. While some readers may be familiar with these fundamental concepts, their inclusion ensures the thesis is both comprehensive and accessible to a broader audience. Readers already well-versed in haematopoiesis may proceed directly to page 4, where the data acquisition and thesis outline are provided.

1.2 Haematopoietic system

Haematopoiesis is the process of blood and immune cell production, primarily occurring in the bone marrow [76]. The haematopoietic system consists of various cell types derived from haematopoietic stem cells (HSCs), with the human body containing approximately 10^{13} mature haematopoietic cells [28, 39, 50, 114, 127].

The haematopoietic process from HSCs to mature haematopoietic cells is typically described by the haematopoietic differentiation hierarchy. In the 21st century and before,



(a): Discrete differentiation model (b): Continuous differentiation model

Figure 1.1: Hierarchical models of haematopoiesis development.

(a) Classic hierarchical models are discrete, where MPP are multipotent progenitors, MGKs are megakaryocytes, DCs are dendritic cells, NK are natural killer cells and ILCs are innate lymphoid cells [25, 82]. (b) Hierarchical models are continuous during the period after 2015 [25, 82].

HSCs were considered a homogeneous population of long-term and short-term HSCs, which could differentiate into common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs) populations [75]. However, with the development of surface markers around 2005 [3], heterogeneity within the HSC population became evident, leading to the classic haematopoietic hierarchy model [25, 82]. In the classical haematopoietic hierarchy, HSCs occupy the top of the hierarchy, characterised by their ability for self-renewal and multipotency [39, 109]. The human body contains $2.5 \times 10^4 - 1.3 \times 10^6$ HSCs, which generate 10^{13} mature haematopoietic cells [28]. Mice have 5×10^3 HSCs, capable of producing 10^{10} haematopoietic cells [28]. HSCs, primarily located in the bone marrow, give rise to multipotent progenitors (MPPs), which serve as key intermediaries in the haematopoietic system [114]. There are approximately 1.4×10^5 MPPs in the mouse [28]. As MPPs differentiate, they lose their multipotency, leading to the production of specific cell types required for haematopoiesis [150].

The classical haematopoietic hierarchy in Figure 1.1 (a) clearly outlines the step-by-step process of haematopoietic stem cell differentiation, presenting it as a discrete model of cellular differentiation. However, single-cell technology challenges the classic haematopoietic hierarchy. This approach reveals the functional and molecular heterogeneity of HSCs, allowing researchers better to explore the relationship between haematopoietic stem and

progenitor cells [14, 19]. As shown in Figure 1.1(b), the differentiation model suggested haematopoietic differentiation is a continuous process rather than a discrete hierarchy [82]. Comparing Figure 1.1 (a) with (b), it is clear that with the development of single-cell technology, the understanding of the process of cell differentiation has changed, such as megakaryocytes (MGKs). MGKs are the largest haematopoietic cells [139], and mouse bone marrow contains 7.7×10^5 MGKs [28]. MGKs are mature haematopoietic cells that strictly regulate the maintenance of HSCs function [18]. Initially, MGKs were thought to be unable to differentiate directly from HSCs, but current research shows that HSCs and MGKs are closer to each other in the continuous haematopoietic differentiation hierarchy than previously recognised [23, 57].

1.3 Haematopoietic microenvironment

Bone marrow is the major site of haematopoiesis, where HSCs differentiate into other haematopoietic cells to maintain the balance of the haematopoietic system [35]. Studies using imaging techniques reveal that HSCs are mainly located near blood vessels in the mouse bone marrow [1, 110, 153]. Haematopoietic stem cells and their surrounding cell types, such as megakaryocytes, mesenchymal stem cells and T cells, constitute the microenvironment, or 'niche' in the bone marrow, which plays a crucial role in regulating the fate and function of HSCs [9, 18, 31, 49, 88]. According to Tobler's first law of geography: 'Everything is related to everything else, but near things are more related than distant things', the spatial balance within the niche is essential for maintaining haematopoiesis [146]. However, changes in the steady-state niche can support the development of leukemogenesis [41]. Leukaemia is a blood cancer. The common symptom of leukaemia patients is cytopenias due to a disruption of normal haematopoiesis [149].

1.4 Leukaemia

Cancer is responsible for approximately 30% of all deaths in Ireland every year, with leukaemia being the most commonly diagnosed cancer in 2021 - 2022 [98]. The most common are acute myeloid leukaemia (AML), and acute lymphocytic leukaemia (ALL) [149]. A better understanding of the dynamics of cancer, particularly AML and ALL, is important. These diseases serve as experimental models to perturb healthy blood production, offering valuable insights into both normal haematopoiesis and disease dynamics.

AML is a malignant haematopoietic stem cell disorder characterised by clonal proliferation of primitive haematopoietic stem progenitor cells [45, 71]. The diagnosis of AML requires identification of 20% or more myeloid blasts by morphological evaluation of pe-

ripheral blood or bone marrow [37]. ALL is a malignant transformation and proliferation of lymphoid progenitor cells in the bone marrow, blood, and extramedullary sites, and its pathogenesis involves abnormal proliferation and differentiation of a clonal population of lymphoid cells [142]. In 2016, the World Health Organisation divided ALL into B-lymphoblastic acute lymphocytic leukaemia (B-ALL) and T-lymphoblastic acute lymphocytic leukaemia (T-ALL) cases, with B-ALL accounting for approximately 75% of cases, while T-ALL accounts for the rest [8].

Leukaemia differs from other solid cancers in that its cells have an innate ability to migrate, allowing them to move rapidly throughout the body over time [149]. The temporal and spatial changes in leukaemia make them a deadly and challenging disease to treat [149]. Haematopoietic stem cell transplantation remains a widely used immunotherapy in clinical practice that treats haematological malignancies like AML and ALL [68, 102, 141]. With the advancement of immunotherapy, chimeric antigen receptor T cells (CAR-T) have been created and shown to significantly improve the treatment outcomes for patients with B-ALL [96, 112].

CAR-T cells, which are genetically modified to target cancer cells, have been approved for clinical use in several countries [33]. The procedure involves drawing the patient's blood, which is then sent to a laboratory where it is modified to create CAR-T cells capable of targeting cancer cells. Once the CAR-T cells are prepared, they are transported back to the hospital and infused into the patient to treat the disease. However, there is limited data on CAR-T therapy efficacy in human solid tumours, making its effectiveness as a therapy uncertain [56].

Despite the success of CAR-T therapy in B-ALL, AML treatment has seen little advancement in over 50 years [113]. Existing studies have shown that T cells, as one of the most important immune cells, play an important role in fighting leukaemia [111, 121]. At the same time, T cell immunotherapy has a significant effect on a variety of solid cancers [138, 156], but the potential of T cells to treat haematopoietic malignancies such as AML remains unclear, so revealing the role of T cells in the tumour microenvironment has become an urgent matter [60].

1.5 Data acquisition and current approaches

Haematopoiesis is a dynamic process, and quantitative analysis of cell movement, differentiation and division can provide a better understanding of the haematopoietic system [28]. This section describes the data acquisition method for the experimental techniques

that were used to generate the data analysed in this thesis, and the current research approaches to analyse these data.

Over the past few decades, intravital microscopy (IVM) has emerged as a popular tool for studying cellular processes *in vivo*, providing real-time single-cell resolution spatio-temporal data within organs [120, 130]. Combining IVM with cell fluorescent labelling has allowed researchers to observe the dynamics of haematopoietic cells in organs [89, 130]. Fluorescent staining enables cells and tissue components to be brightly labelled, allowing for simultaneous visualisation of multiple cell types *in vivo* [120]. There are three primary methods for fluorescent cell labelling. The first is to genetically modify the experimental animals to introduce one or more fluorescent proteins into their genome [105]. The second method is antibody staining after sampling, which allows for the detection of cell structures and types within the tissue [2]. The third approach uses viral vectors to deliver genetic tools into cells that are adoptively transferred into secondary animals [65]. For example, lentiviral vectors can introduce fluorescent proteins, such as the mScarlet gene, which is replicated through reverse transcription, integrating the gene into the cell's DNA.

IVM techniques can be applied to different organs, such as cremaster muscle, lymph nodes, liver, and bone marrow [67, 89, 126, 147]. IVM of murine bone marrow has greatly promoted the understanding of the haematopoietic process of transplanted haematopoietic stem and progenitor cells in the bone marrow [27, 89]. IVM is also often used to study the spatio-temporal movement of cells in the mouse calvarium bone marrow, with tracking time ranging from days to weeks [40]. IVM can also image long bones to reveal the spatial distribution of cells at a fixed time [73]. IVM combined with fluorescent staining provides one of the most accurate methods to visualise single cells over fixed or short periods of time, significantly enhancing our understanding of haematopoiesis [26]. This approach generates large, high-dimensional, and complex image datasets containing real-time spatio-temporal biological information. However, as these datasets grow in complexity, more advanced spatio-temporal analysis techniques are necessary to fully understand the data. For example, while current image analysis techniques have successfully extracted haematopoietic process information in mouse bone marrow [52, 74, 103, 152], further refinement using more sophisticated spatio-temporal statistical methods could offer deeper insights.

When using IVM to visualise the spatial distribution of cells in bone marrow at a fixed time, researchers typically capture 35-40 μm thick slices of bone marrow, which limits detection to a subset of cells [152]. Since cell distribution in the bone marrow may not

be random, analysing only a partial view may overlook important spatial relationships and lead to incomplete insights. While this approach is widely used, adopting advanced imaging techniques could provide a more comprehensive picture of cell distribution and improve analysis. Additionally, current studies often overlook the importance of spatial statistical processes and frequently use a spatial analysis technique called random dots, which tests whether cell distribution is uniform within a given space [30, 137, 152]. While this approach is intuitive, it focuses on determining cell distribution throughout the space, and does not quantify the spatial relationships between multiple cell populations. As a result, spatial interactions between various types of cells may be ignored. By incorporating more advanced spatial statistical methods, researchers can better understand not only the distribution but also the spatial relationships between different cell types, providing a deeper biological understanding.

While significant progress has been made in three-dimensional data extraction and analysis, there is still room for improvement in the analysis of four-dimensional, spatio-temporal, data obtained from IVM [52, 53, 74, 103, 152]. Spatio-temporal statistical analysis would allow researchers to better track dynamic changes in the bone marrow over time, rather than relying on fixed time points. The main difficulty with high-dimensional spatio-temporal data lies in effective extraction and interpretation. Current software tools such as '`Imaris`' (commercial) and '`ImageJ`' (open-source) could provide basic spatio-temporal data extraction and analysis [90, 103]. However, combining these softwares with more robust statistical frameworks would significantly improve the accuracy and reliability of spatio-temporal data extraction and interpretation. Integrating advanced spatial and temporal statistical pipelines with *in vivo* studies would allow for a more effective investigation of dynamic cellular behaviours, further advancing our understanding of haematopoiesis.

Lastly, the human body contains approximately 10^{13} haematopoietic cells. Without accounting for cell death, HSCs would require at least 44 divisions to expand and generate a sufficient number of haematopoietic cells [28]. The blood system undergoes continuous change, with cells constantly growing, dividing, and differentiating. While IVM is a powerful tool for observing single-cell behaviour over short time frames, challenges remain for longer-term studies. Cells often move out of the field of view, and the current IVM technique can only track up to eight generations, falling short of the Hayflick limit of 40 to 60 generations [63]. Developing an advanced method for long-term cell tracking and robust temporal statistical analyses is essential for capturing the full scope of cell dynamics.

1.6 My contribution

As discussed earlier, traditional statistical methods lack a robust statistical framework for analysing and interpreting complex high-dimensional data across different time scales and space scales. This research addresses these limitations by integrating advanced spatial and temporal statistical techniques with data from collaborators. By introducing practical, reproducible, and generalizable statistical methods, this thesis presents a novel framework for tackling the challenges of high-dimensional spatio-temporal data analysis.

My contributions are divided into three parts based on different time scales. At fixed time points, a spatial statistical pipeline was developed to quantify the distribution of various cell types and assess their spatial relationships, capturing complex biological information. For a time around one day, a statistical framework was designed to integrate data-cleaning processes, and analyse the relationship between cell movement and differentiation, revealing dynamic patterns. For long-term periods, mathematical models and statistical methods were applied to estimate the average number of cell divisions *in vivo*, providing insights into long-term cell generational distributions.

By combining advanced spatial and temporal statistical methods, this framework surpasses traditional approaches by capturing the spatio-temporal relationships between different cell types. The statistical methods presented in this thesis offer a deeper understanding of the dynamic processes underlying cellular behaviour.

1.7 Thesis outline

This thesis focuses on analysing the spatio-temporal distribution of cells to better understand cellular fate over time. It is divided into three main chapters to address different data types and timescales: spatial distribution of cells at a fixed time, cell movement and differentiation over a day, and long-term cell division patterns.

In Chapter 2, a spatial statistical pipeline is introduced and implemented, based on a dataset of millions of fluorescently stained cells in bone marrow at a fixed time. The chapter begins by analysing a single cell type and progresses to automatically identify areas of high cell density, while also assessing relationships between multiple cell types. Non-parametric tests are applied, and 3D geographically weighted models are created to understand the spatial distribution of cells at a fixed time, providing a detailed snapshot of their three-dimensional arrangement. The work from this chapter has an associated preprint [2] and is submitted for publication.

Chapter 3 introduces a statistical framework for cleaning high-dimensional spatio-temporal data derived from time-lapse microscopy, and applies unsupervised machine learning methods to study motion-based cell differentiation. This chapter focuses on the dynamic behaviour of early engrafted haematopoietic stem and progenitor cells and how these cells evolve around a day-long period. This work is currently being prepared for submission.

Chapter 4 presents techniques for analysing the generational distribution of cells over a one-day period, providing insights into the evolution of cell populations over time. This chapter introduces a novel method called DivisionCounter, developed in collaboration with Dr. Leïla Perié's lab, for analysing long-term generational cell distributions *in vivo*. The statistical framework used in this chapter is based on data collected by collaborators, and the work has an associated preprint [65] and submitted for publication.

The details of the experiments used for analysis were described in Appendix A for self-contained and completeness. The experiment details and 3D data extractions for statistical analysis of spatial distribution in large cell populations presented in Chapter 2 were shown in Section A.1, performed by our collaborators in Prof. Cristina Lo Celso's lab at Imperial College London. Section A.2 showed the biological experimental details in Chapter 3 for tracking the movement of haematopoietic stem and progenitor cells using intravital microscopy, which was executed by our collaborators in Prof. Cristina Lo Celso's lab. The experiments for Chapter 4 about cell generational distribution conducted by Dr. Leïla Perié's lab at Curie Institute were presented in Section A.3.

2

Statistical analysis of spatial distribution in large cell populations

2.1 Abstract

This chapter presents the spatial analysis of work 'PACESS: A practical AI-based cell extraction and spatial statistical pipeline for large bone marrow tissue images' accomplished in collaboration with partners in Prof. Cristina Lo Celso's lab at Imperial College London, a version of which has been made available as a pre-print [2]. The main contributions of this work are that, from large, complex 3D images of bone marrow, the three-dimensional spatial position of each cell was obtained by using two-dimensional cell annotation, and the relationship between different cell types was analysed by using spatial statistics. My main contribution was to design and implement a spatial statistical analysis pipeline to analyse the spatial distribution of cells at a fixed time.

A central tenet of biology is that patterns of cellular relationships correspond to function. Multi-colour fluorescent staining is a rapidly advancing technique in molecular biology that is being used to locate and classify hundreds of thousands of individual cells within tissues, providing massively improved observational resolution. Appropriate statistical methods are needed to quantitatively and confidently draw inferences on complex multicellular interactions from such large, three-dimensional spatial datasets. Using, as an exemplar, a primary dataset consisting of millions of cells fluorescently stained cells in murine bone marrow is employed to report on an approach to achieve these goals. The spatial analysis involves generating exploratory statistics that quantitatively assess spatial heterogeneity in individual cell populations; identifying regions of abnormally high cellular density; performing hypothesis tests to determine the locality-dependent influence of one cell type's density on another; and, finally, formulating a 3D geographically weighted regression approach that quantifies the location-dependent relationships between cell types.

This innovative and robust spatial analysis pipeline holds broad potential for unravelling complex multicellular interactions and unlocking the intricacies of tissue complexity.

2.2 Introduction

The cellular architecture of organs is known to be a key feature for the maintenance of their function, and changes in cellular organisation are often linked to disease. This has been extensively demonstrated for tissues that have well-known structures such as the skin or the brain [48, 104]. The bone marrow is the organ where haematopoiesis, the process of maintaining all blood cells' turnover through the daily generation of billions of diverse cell types in mice, and trillions in humans, takes place [28]. The combination of flow cytometry, single-cell transcriptomics and functional assays has provided a wealth of information on the different haematopoietic cell populations co-existing in the bone marrow [25, 80, 82].

However, the spatial organisation of this tissue remains poorly understood. Bone marrow is the semi-solid tissue in the body and has long been considered an amorphous, even liquid, structure with a wide range of different cell types tightly packed within a confined cavity. These features make the bone marrow a uniquely challenging tissue for histological imaging and quantification [29, 85, 137]. To address this challenge, researchers in the field have made use of increasingly sophisticated bone marrow multidimensional imaging techniques, 3D or greater, focusing on visualising haematopoietic stem and progenitor cell (HSPC) populations in large bone marrow tissue preparations [29, 110, 157]. While these led to the emerging concept that the bone marrow is spatially and regionally organised to differentially support distinct HSPC populations, the spatial organisation of multiple hematopoietic cell types remains unclear [4, 157]. Understanding it holds clues for the more effective harnessing of HSPC function, development of improved therapies for haematological disorders and overall healthier ageing.

Quantification of multicellular interactions within large bone marrow 3D image datasets has not yet been achieved for two main reasons. First, image segmentation by intensity threshold, which is, historically, the most widely used method for identifying cells of interest, underperforms in scenarios where cells are compressed or tightly packed together, as in the bone marrow, where boundaries between cells become more difficult to distinguish. Hence, a high number of cells are lost in the process. The second challenge relates to the statistical methodologies available to assess multicellular interactions. Inferences have typically been made using pairwise hypothesis testing, but this approach is limited as it does not facilitate the direct examination of the relationships between a higher number of

cell types, making it challenging to formulate definitive conclusions from the underlying data.

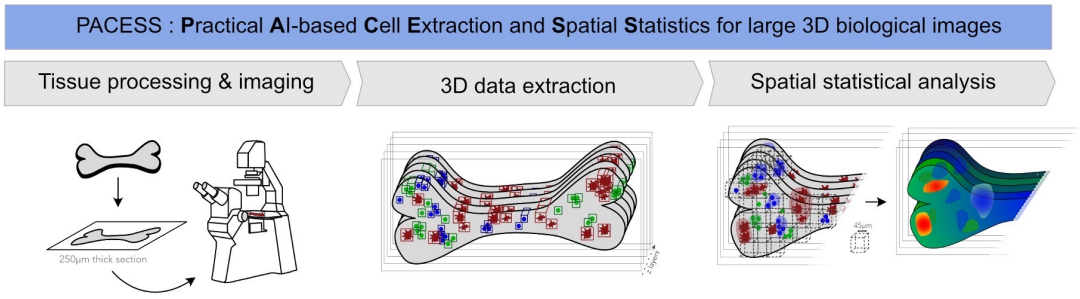


Figure 2.1: PACESS overview.

The workflow consists of three steps: 1) tissue processing and imaging (*in vivo* and *ex vivo*), 2) 3D data extraction using object detection neural network trained on 2D data and scale to 3D, and 3) spatial statistical analysis consisting in spatial inhomogeneity quantification, automatic identification of areas of high cellular density and statistical assessment of dependencies between cell types.

PACESS (Practical AI-based Cell Extraction and Spatial Statistics), as presented in Figure 2.1, is a workflow for extracting and analysing haematopoietic cells from large 3D bone marrow images, overcoming existing limitations. This workflow makes use of convolutional neural network object detection to classify and identify the locations of cells in 3D. An augmented object detection deep neural network trained using 2D data alone is introduced, for which images can be rapidly annotated. Annotations are created for each image layer in the 3D data and the output from multiple layers is automatically combined to identify each cell's type and location within the 3D space. Once the spatial data are extracted, a set of 3D spatial statistical analyses are applied. The steps in this process include: generating exploratory statistics that quantitatively assess spatial heterogeneity; identifying regions of high cellular density of specific cell types; and, finally, identification and visualisation of location-dependent relative abundance between cell types. The resulting analysis is presented as a holistic 'statistical map' of bone marrow tissue.

PACESS is applied to study the spatial interactions of megakaryocytes (MGKs) and T cells in the context of leukaemia. Leukaemia progression leads to healthy haematopoietic cells being dislodged [4, 119], and therefore represents a good model for studying changes in multiple cell populations' spatial interactions. First, PACESS was able to efficiently extract the cellular information for the three cell populations, even in areas tightly packed with leukaemia. Then, the statistical analysis revealed that the density of AML cells at a

specific bone marrow location will significantly affect T cells and MGKs densities, albeit with unique patterns, resulting in the generation of multiple microenvironments characterised by unique relative densities of these three cell types.

By its nature, the bone marrow is an organ where cells are highly packed together. This is exacerbated during acute myeloid leukaemia (AML) when malignant cells outcompete healthy haematopoiesis through mechanisms that are still not fully understood. As an illustrative example, our biological collaborators at Imperial College London selected a $2755 \times 5805 \times 155\ \mu m^3$ AML cells infiltrated bone marrow sample, which was sectioned with a distance of 5 μm between each slice. Then, using an object detection deep neural network trained on 2D data alone, for which images can be rapidly annotated, annotations are created for each image layer in the 3D data and the output from multiple layers is automatically combined to identify each cell's location and cell-type within the 3D space.

The 3D data extraction process was primarily developed by a collaborator in Prof. Lo Celso's lab, and is not included in this thesis, as detailed in [2]. The neural network was trained using 18,240 manual annotated cell examples. The data collection process identified the classification and location of AML cells, T cells, and MGKs. Cells were the bone marrow identified as being outside excluded from the object detection dataset. Finally, the spatial analysis was performed on a dataset consisting of 163,683 AML, 8,308 T cells, and 2,191 MGKs.

2.3 Methods

This section describes the spatial statistical pipeline developed to analyse the complex 3D data obtained from the extraction step, as shown in Table 2.1. To sequentially address biologically significant questions, the pipeline begins by exploring the spatial distribution of a single cell type, progresses to analysing the spatial interactions between two cell types, and finally, accesses the relationships among multiple cell types. All programming was performed using `R` (version `R` 4.3.2).

Variable	**Detail**	**Description**
x	Numerical variable	Three-dimensional coordinates
y	Numerical variable	Three-dimensional coordinates
z	Numerical variable	Three-dimensional coordinates
Cell type	Character	AML, T cell, MGK

Table 2.1: Variables in the original dataset.

2.3.1 Spatial inhomogeneity quantification

The first biologically relevant question was to quantify the spatial distribution of a given cell type within the bone marrow. That is, to determine if it was spread out, randomly distributed, or significantly. To that end, Moran's I index was used.

The positions and classifications of cells were obtained from the 3D data extraction process. Cells outside the bone marrow were excluded from the object detection dataset. Then, the bone marrow was divided into $\kappa(\mu m)^3$ cubes. The midpoint of the kth cube $L_k = (x_k, y_k, z_k)$ was used as the three-dimensional geographic coordinates, and the number of cells of each type in each cube was recorded. The data in each of these cubes was used as input to the analysis and model.

Moran's I index was used to measure spatial autocorrelation for each cell type [44]. The formula for Moran's I index is

$$I = \frac{n \sum \sum_{i \neq j} \varphi_{ij} (\omega_i^{\tau_k} - \overline{\omega^{\tau_k}})(\omega_j^{\tau_k} - \overline{\omega^{\tau_k}})}{(\sum_{i=1}^{n} (\omega_i^{\tau_k} - \overline{\omega^{\tau_k}})^2) \, (\sum \sum_{i \neq j} \varphi_{ij})},$$

where n is the total number of cubes, $\omega_i^{\tau_k}$ is the number of cells in type k at the ith cube, $\omega_j^{\tau_k}$ is the number of cells at the jth cube, $1 \leq i, j \leq n$, $i, j \in \mathbb{N}$, $\overline{\omega^{\tau_k}}$ is the mean of the number of type k cells at each cube, and φ_{ij} is a spatial weight. The formula for φ_{ij} is

$$\varphi_{ij} = \begin{cases} 1 & \text{if } d_{ij} \leq \kappa \\ 0 & \text{if } d_{ij} > \kappa \end{cases},$$

where d_{ij} be the Euclidean distance between the centroids of cube i and cube j

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \, .$$

The range of Moran's I is $[-1, 1]$. If Moran's I is positive, cells tend to be aggregated in common areas. The yellow cubes in Figure 2.2 have similar cell numbers to the red cube. If Moran's I is close to zero, cells are distributed randomly in space. When its value is less than zero, cells are more homogeneously dispersed than one would expect from a random process. The yellow cubes in Figure 2.2 have different cell numbers to the red cube.

2.3.2 Automatic identification of areas based on cellular density

For cell types for which Moran's I indicated suggested significant aggregation, a method was required to automatically identify 3D regions of high density. For that purpose,

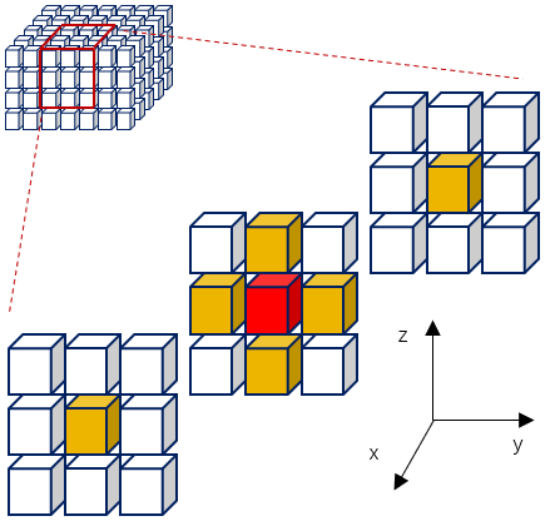


Figure 2.2: Schematic representation of Moran's I.

The neighbours of the red cube are immediately adjacent yellow ones that have a face in common.

the density-based spatial clustering of applications with noise (DBSCAN) algorithm was employed to cluster the cells [129].

For cube (x_i, y_i, z_i), $\boldsymbol{N}_i = \{(x_j, y_j, z_j) | d_{ij} \leq \kappa\}$ is a set of all neighbouring cubes that are κ μm or less away from the ith cube. The number of cells in $\boldsymbol{N}_i$ is recorded as $\|\boldsymbol{N}_i\|$. When the faces of cubes are connected to each other, they are neighbours, so each cube in this case has six neighbours, which has been shown in Figure 2.3.

High-density area identification

If the total number of cells in cube (x_i, y_i, z_i) and its neighbours is greater than 7, the number of cubes being considered, times γ^h, such as the third quartile of counts, then the ith cube and its neighbours are marked as high-density cubes.

Let $\Omega^h = \left\{(x_j, y_j, z_j) | \, \|\boldsymbol{N}_i\| \geq 7\gamma^h, d_{ij} \leq \kappa\right\}$ is the set which includes all high-density cubes. For any cube $(x_i, y_i, z_i) \in \Omega^h$, $\partial_1^h = \left\{(x_j, y_j, z_j) | d((x_i, y_i, z_i), \Omega^h) \leq \kappa\right\}$ where ∂_1^h includes all high-density cubes close to ith cube, and $d(A, B)$ represents Euclidean distance between the set A and the set B. Then, $\partial_2^h = \left\{(x_j, y_j, z_j) | d(\partial_1^h, \Omega^h) \leq \kappa\right\}$, $\cdots$, $\partial_{n+1}^h = \left\{(x_j, y_j, z_j) | d(\partial_n^h, \Omega^h) \leq \kappa\right\}$. When $\left|\partial_{n+1}^h\right| - \left|\partial_n^h\right| = 0$ where $\left|\partial_n^h\right|$ is the number of cubes in the ∂_n^h, iteration ends and ∂_n^h is the first cluster which is records as C_1^h. If $\Omega_1^h = \Omega^h - C_1^h = \varnothing$, then there is one cluster. Conversely, any cube $(x_l, y_l, z_l) \in \Omega_1^h$ are selected. The second cluster C_2^h and Ω_2^h can be obtained using the same step. When

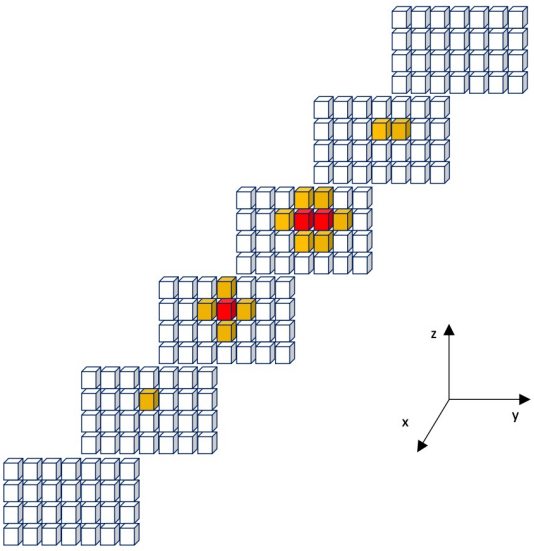


Figure 2.3: Schematic representation of DBSCAN algorithm.

The neighbours of the red cube are adjacent yellow cubes. Each cube has six neighbours. In this pictorial representation, there are three red cubes surrounded by yellow cubes that are all part of the cluster.

$\Omega^h_{p+1} = \varnothing, p \in \mathbb{N}^*$, the data has p clusters. In addition, for any cube that does not belong to any cluster, these cubes are in the set C^h_0.

2.3.3 Statistical assessment of dependencies between cell types

Spatial permutation tests

Having identified regions of unusually high density in one cell type, a natural biological question is whether it influences the distribution of other cell types. To statistically challenge the hypothesis that there is no effect, a permutation test approach was adopted. In detail, after cell clustering, the information on all clusters can also be obtained. For the ith cluster ($i \in \mathbb{N}^*$ and $i \leq p$), $U_i = \{(x_j, y_j, z_j) | 0 < d((x_j, y_j, z_j), C_i) \leq \kappa\}$ is a set which contains the cubes in the ith cluster and the cubes κ μm away from the ith cluster, and these cubes do not belong to any other clusters. Hence, the cubes around the ith cluster are $B_i = \{(x_j, y_j, z_j) | (x_j, y_j, z_j) \in U_i, (x_j, y_j, z_j) \notin C_i\}$.

Permutation tests were used to detect changes in the number of cells in the cluster as well as the number of cells around the cluster [83]. The null hypothesis for the permutation test is that the mean number of cells in the cubes is independent of whether the cubes are in C_i or B_i. Here, $A = (\omega_{i_1}, \omega_{i_2}, \cdots, \omega_{i_{|C_i|}}, \omega_{i_{|C_i|+1}}, \omega_{i_{|C_i|+2}}, \cdots, \omega_{i_{|C_i|+|B_i|}})$ is an ordered observations set, where ω_{i_k} is the number of cells in the kth cube in the C_i. In the set A,

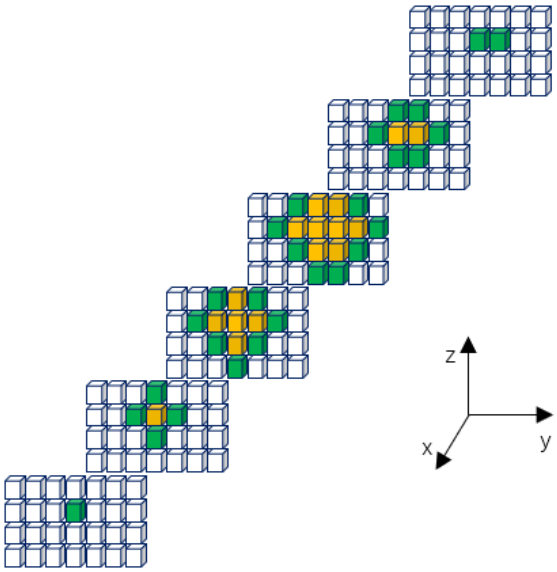


Figure 2.4: Schematic representation of spatial permutation test.

The yellow cubes are all part of one cluster. The green cubes indicate the boxes that are in the boundary region surrounding the cluster.

the first $|C_i|$ elements are the number of cells in the ith cluster, and the last $|B_i|$ elements are the number of cells around the cluster. Hence, the real-valued statistic is used under this null hypothesis

$$M(A) = \left| \frac{\sum_{i=1}^{|C_i|} \omega_i}{|C_i|} - \frac{\sum_{j=|C_i|+1}^{|C_i|+|B_i|} \omega_j}{|B_i|} \right|.$$

Then, a permutation π is created that reassigns labels to individual data. A reordered observation set is obtained:

$$A_\pi = (\omega_{\pi(i_1)}, \omega_{\pi(i_2)}, \cdots, \omega_{i_{\pi(|C_i|)}}, \omega_{i_{\pi(|C_i|+1)}}, \omega_{i_{\pi(|C_i|+2)}}, \cdots, \omega_{i_{\pi(|C_i|+|B_i|)}}).$$

Similarly, a new reordered set also generates statistics $M(A_\pi)$. A collection of permutations Q can be characterised so that reorderings $\{A_\pi\}_{\pi \in Q}$ are equally likely under the null hypothesis. Then, the empirical distribution of $M(A_\pi)_{\pi \in Q}$ is used to compare with $M(A)$. Therefore, the Monte Carlo approximation with 500,000 test statistics is applied to estimate the two-tailed p-value, and the Bonferroni correction is performed to avoid Type I errors [83].

Logistic regression

To statistically model the extent of the influence of one cell type's presence over another, the logistic regression was performed to assess whether the presence or absence of other cell types was independent of dense areas, which has been shown in Figure 2.5.

The formula for binary logistic regression is:

$$\text{logit}\left(P(\Omega^h)\right) = \beta_0 + \sum_{k=1}^{m} \beta_k e_k + +\epsilon \ ,$$

where, $P(\Omega^h)$ is the probability of the cube in cluster areas, and $\text{logit}\left(P(\Omega^h)\right) = \log\left(\frac{P(\Omega^h)}{1-P(\Omega^h)}\right)$. The β_0 is the intercept in the model, m is the number of explanatory variables, β_k is the coefficient for the kth input variable, e_k is the kth explanatory variable, and ϵ is the error term of the model.

Logistic regression is used to predict the likelihood of the presence of cell types in a cluster versus cells not in a cluster. Each cube is changed into a dummy variable as the response variable based on whether it is in the cluster or not. The explanatory variables are binary variables, also determined by the presence or absence of the cell type. The statistical significance of each binary variable is assessed using t-tests. The logistic regression was performed by the '`glm`' function in '`R`' (version 4.4.0).

Three dimensional geographically weighted regression (3D GWR)

While the methods used so far have largely been for single cell types or pairs of cell types, the geographically weighted regression model (GWR) allows a more holistic analysis of all cell types simultaneously, albeit with a more involved interpretation of the model results.

GWR was used to examine the spatial relationship between explanatory variables and the response variable [47]. The standard GWR model is defined for two-dimensional (2D) planes, and this work extended to three-dimensional (3D) space. The natural extension of the 2D GWR approach to 3D is

$$r(x_i, y_i, z_i) = \beta_0(x_i, y_i, z_i) + \sum_{k=1}^{m} \beta_k(x_i, y_i, z_i) e_k(x_i, y_i, z_i) + \epsilon(x_i, y_i, z_i) \ ,$$

where $r(x_i, y_i, z_i)$ is the response variable at the ith cube, $\beta_0(x_i, y_i, z_i)$ is the intercept in

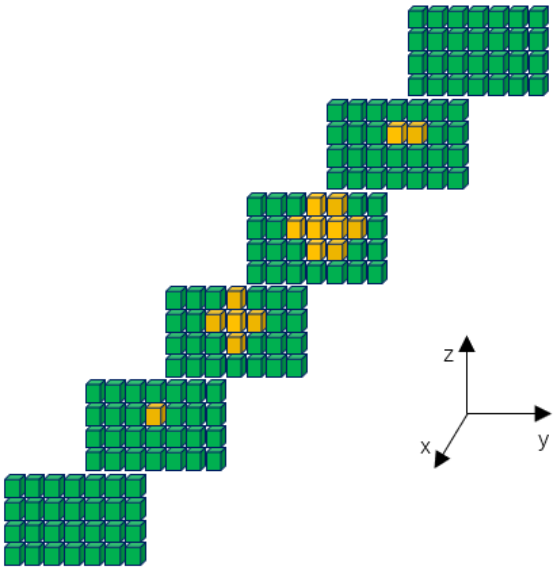


Figure 2.5: Schematic representation of logistic regression.

The yellow cubes are all part of one cluster. The green cubes indicate the boxes that are out of the cluster.

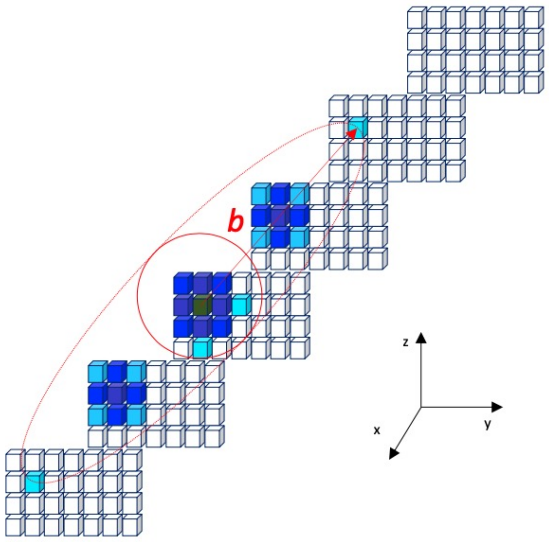


Figure 2.6: Schematic representation of 3D geographically weighted regression.

Three-dimensional geographically weighted regression quantifies the spatial relationship between cell counts of cells of different types. The red arrow represents the bandwidth range. The blue shading shows the influence of the centre cube on each cube within the bandwidth.

the model, m is the number of explanatory variables, $\beta_k(x_i, y_i, z_i)$ is the coefficient for the kth input variable at the ith cube and $e_k(x_i, y_i, z_i)$ is the kth explanatory variable at the ith cube. In addition, $\epsilon(x_i, y_i, z_i)$ is the error term at the ith cube.

A weighted least squares method is used to get the coefficients $\hat{\beta}(x_i, y_i, z_i) = (\beta_0(x_i, y_i, z_i), \beta_1(x_i, y_i, z_i), \cdots, \beta_m(x_i, y_i, z_i))^T$. The formula for the coefficients is

$$\hat{\beta}(x_i, y_i, z_i) = \left(E^T W(x_i, y_i, z_i) E\right)^{-1} E^T W(x_i, y_i, z_i) r \ ,$$

where

$$E = \begin{pmatrix} 1 & e_1((x_1, y_1, z_1) & e_2(x_1, y_1, z_1) & \cdots & e_m(x_1, y_1, z_1) \\ 1 & e_1(x_2, y_2, z_2) & e_2(x_2, y_2, z_2) & \cdots & e_m(x_2, y_2, z_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & e_1(x_n, y_n, z_n) & e_2(x_n, y_n, z_n) & \cdots & e_m(x_n, y_n, z_n) \end{pmatrix} \in \mathbb{R}^{n \times (m+1)} \ ,$$

$$r = \begin{pmatrix} r(x_1, y_1, z_1) \\ r(x_2, y_2, z_2) \\ \vdots \\ r(x_n, y_n, z_n) \end{pmatrix} \in \mathbb{R}^{n \times 1} \ ,$$

$$W(x_i, y_i, z_i) = \begin{pmatrix} w_1(x_i, y_i, z_i) & 0 & \cdots & 0 \\ 0 & w_2(x_i, y_i, z_i) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & w_n(x_i, y_i, z_i) \end{pmatrix} \in \mathbb{R}^{n \times n} \ .$$

The $W(x_i, y_i, z_i)$ is the diagonal weighted matrix at position (x_i, y_i, z_i), and it is determined by the distance between cubes and the kernel function. In this study, the bi-square kernel function is used:

$$w_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b}\right)^2\right)^2 & \text{if } |d_{ij}| < b \\ 0 & \text{otherwise} \end{cases} \ ,$$

where b is the bandwidth. The bi-square kernel function reflects that neighbouring points have more influence on the ith cube than distant ones. Its scaling is determined by a

bandwidth b that is selected by minimising a corrected version of Akaike Information Criterion (AICc) [51]:

$$\mathrm{AICc}(b) = 2n\ln(\hat{\sigma}) + n\ln(2\pi) + n\frac{n+\mathrm{tr}(S)}{n-2-\mathrm{tr}(S)} ,$$

where, $\hat{\sigma}$ is the standard deviation of the residuals, and $\mathrm{tr}(S)$ is the trace of the matrix S, which is called the hat matrix in standard GWR. In detail, the hat matrix is the projection matrix from the observed $r(x_i, y_i, z_i)$ to the fitted value $\hat{r}(x_i, y_i, z_i)$, so the matrix S is

$$S = \begin{pmatrix} E_1^T(E^T W(x_1,y_1,z_1)E)^{-1}E^T W(x_1,y_1,z_1) \\ E_2^T(E^T W(x_2,y_2,z_2)E)^{-1}E^T W(x_2,y_2,z_2) \\ \cdots \\ E_n^T(E^T W(x_n,y_n,z_n)E)^{-1}E^T W(x_n,y_n,z_n) \end{pmatrix} \in \mathbb{R}^{n\times n} ,$$

where, $E_i^T = (1\ e_1(x_i,y_i,z_i)\ e_2(x_i,y_i,z_i)\ \cdots\ e_m(x_i,y_i,z_i)) \in \mathbb{R}^{1\times(m+1)}$ ([91]).

Adapting principles from 2D GWR diagnostics, for the 3D GWR the following diagnostic statistic was adopted, the local R^2. Local R^2 can reflect the quality of local models to explain local data. Local R^2 is defined as [47]

$$R^2(x_i,y_i,z_i) = 1 - \frac{\sum_{j=1}^n w_i(x_j,y_j,z_j)(r(x_j,y_j,z_j) - \hat{r}(x_j,y_j,z_j))^2}{\sum_{j=1}^n w_i(x_j,y_j,z_j)(r(x_j,y_j,z_j) - \bar{r})^2},$$

where, $\bar{r}$ is the mean of response variables, and $\hat{r}(u_i,v_i,z_i)$ is the fitted $r(u_i,v_i,z_i)$. In addition, $w_i(x_j,y_j,z_j)$ is from the weighted matrix.

2.4 Results

The first step, tissue processing and imaging, and the second step, 3D data extraction, in the PACESS workflow were designed and carried out by our collaborators in Prof. Lo Celso's lab. In this section, we demonstrate the spatial statistical analysis pipeline using a bone marrow sample of 2755 × 5805 × 155 μm^3, which includes 163,683 AML cells, 8,308 T cells, and 2,191 MGKs.

In order to visualise and analyse the spatial distribution and relationships between cell

types, space was discretised into non-overlapping, adjacent cubes that cover the entire 3D area. The number of cells of each type in each cube was recorded. For meaningful visualisation, the discretisation needs to be sufficiently coarse that some aggregation of cell counts occurs. A cube size of 45 $(\mu m)^3$ was selected to be sufficiently fine that geographic resolution was retained, but sufficiently coarse that the resulting data could still be computationally assessed without undue burden. With that discretisation, the density of AML cells, T cells and MGKs as a function of their position for a selection of z-depths could be reported in Figure 2.7.

2.4.1 Spatial heterogeneity

AML cells were observed to be more prevalent than the other two types of cells, with most of them appearing to be located in a single mass, as previously seen in Figure 2.7. This observation highlighted the importance of being able to quantitatively assess the homogeneity of the spatial distribution for each cell type, which could be achieved through the calculation of a statistic called Moran's I.

Moran's I for the AML cells was 0.81, quantitatively substantiating the observation that AML cells were largely concentrated in relatively few patches. For T cells and MGKs, respectively, approximately 70% and 90% of cubes recorded a zero cell count, consistent with these cells being less abundant than AML cells. Moran's I of 0.05 for MGKs indicated these cells had a weak geographical dependency resembling random locations. Moran's I for T cells was 0.34, suggesting some positive spatial clustering but less than found for AML cells.

2.4.2 Automatic identification of areas of high cellular density

While Moran's I indicates that AML cells are largely co-located, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm is used to identify the regions of high density [129]. For these data, when the distance between the cubes was $45\mu m$ or less they were considered to be neighbours, resulting in each cube having seven neighbouring cubes that have a face in common (Figure 2.2). If the average number of cells in a cube and its neighbours was more than the third quartile for a single cube, where only a quarter of cubes have more cells than this value, it was considered a dense neighbourhood. Dense neighbourhoods were agglomerated using the DBSCAN algorithm to form contiguous spatial clusters in Figure 2.3.

For the AML data, DBSCAN identified 18 distinct clusters, the largest three of which accounted for 61.8%, 5.9% and 2% of all AML cells. These clusters are shown in Figure

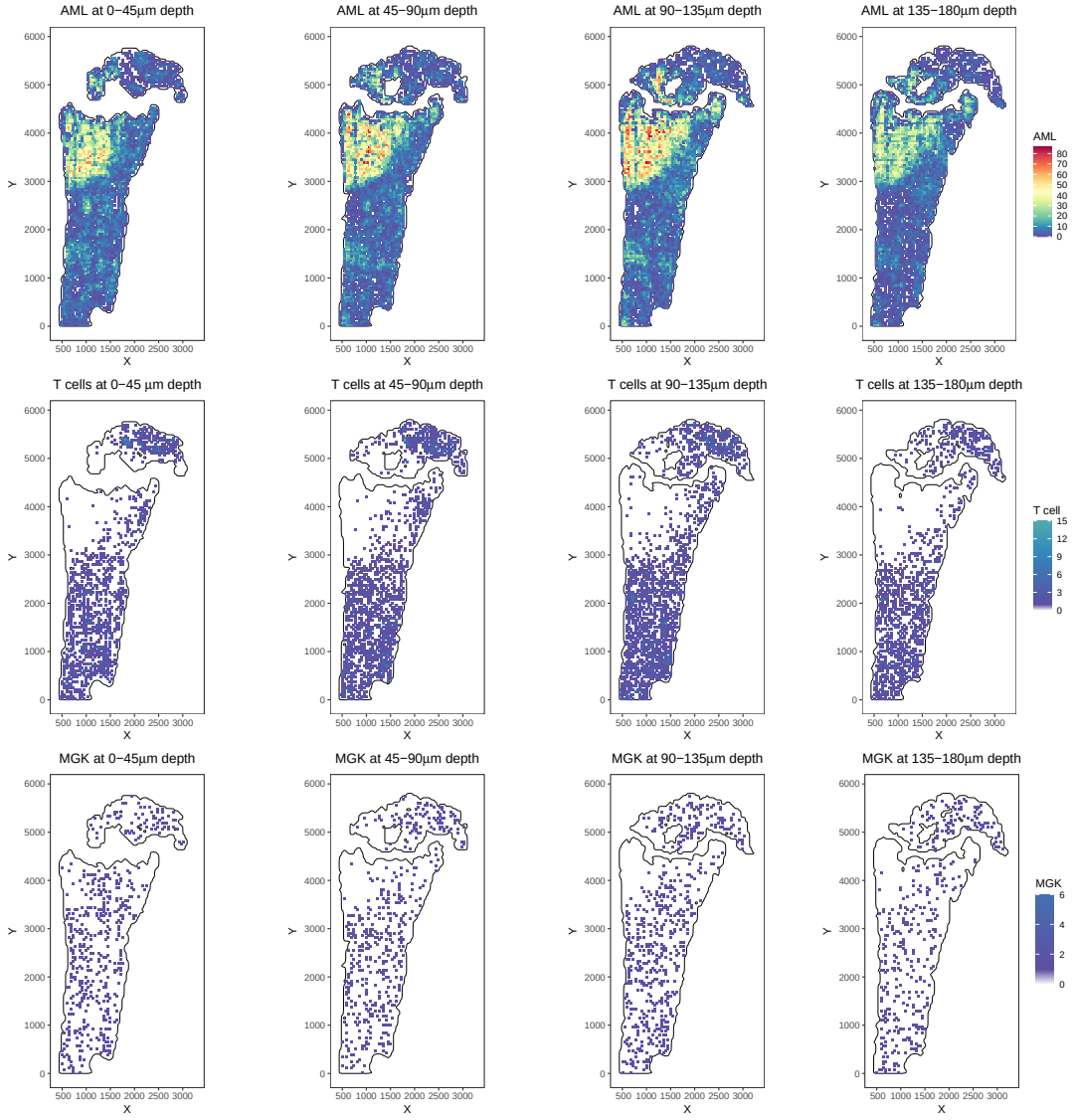


Figure 2.7: Quantitative 2D projection visualisation.

Quantitative Data are aggregated into $45\mu m^3$ cubes, at a range of z-depths heat maps showing the coordinates and the number of AML cells, T cells and MGKs. The colour scale indicates the density of cells. The black line represents the boundary of the bone marrow.

2.8, where these clusters by cell count are marked in decreasing order. The two largest clusters of AML cells, 1 and 2, were proximal to the growth plate, on either of its two sides. Cluster 3 was more distal, and overall the distal region of bone marrow contained much smaller clusters. Interestingly, larger clusters were located adjacent to the endosteum which is a thin vascular membrane lining the cavity of long bones. While centrally located clusters tended to be smaller. As mentioned before, for T cells and MGKs, more than 70% of the cubes recorded zero cell counts, so DBSCAN did not return any clustered regions. This was consistent with these cells being less abundant than AML cells, generally not clustered and returning Moran's I values that were positive but close to zero.

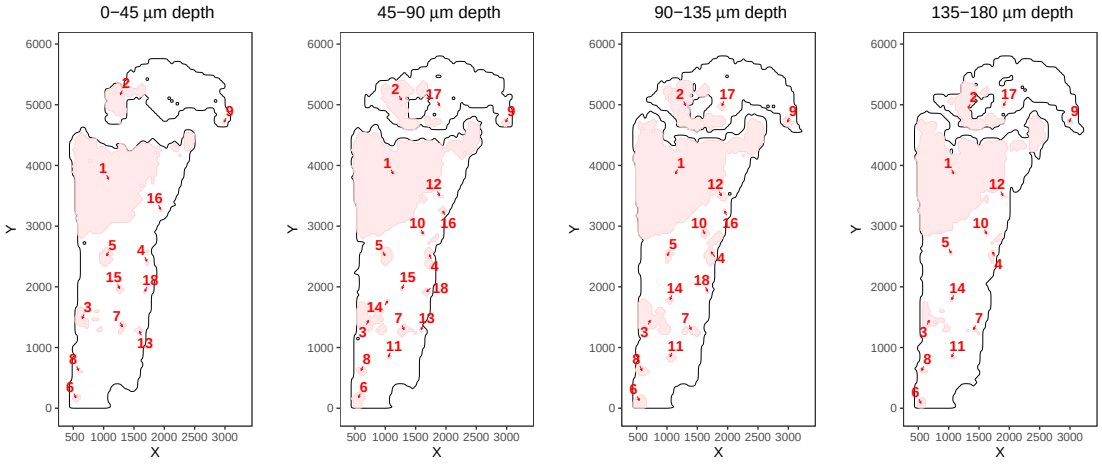


Figure 2.8: AML cells cluster identification.

Two-dimensional projection visualisations at different z-depths show the AML clusters identified using DBSCAN. The clusters are numbers in decreasing order of total AML cell numbers. The black line represents the boundary of the bone marrow.

2.4.3 Spatial density dependencies

The key to understanding the principles regulating multicellular tissues is based on understanding how diverse cell types interact with each other to enable tissue function. It is therefore important to explore whether local heterogeneities in the distribution of specific cell types depend on those of other cell types and to identify them. Permutation tests, logistic regression and 3D geographically weighted regression were used to investigate this.

Spatial permutation tests

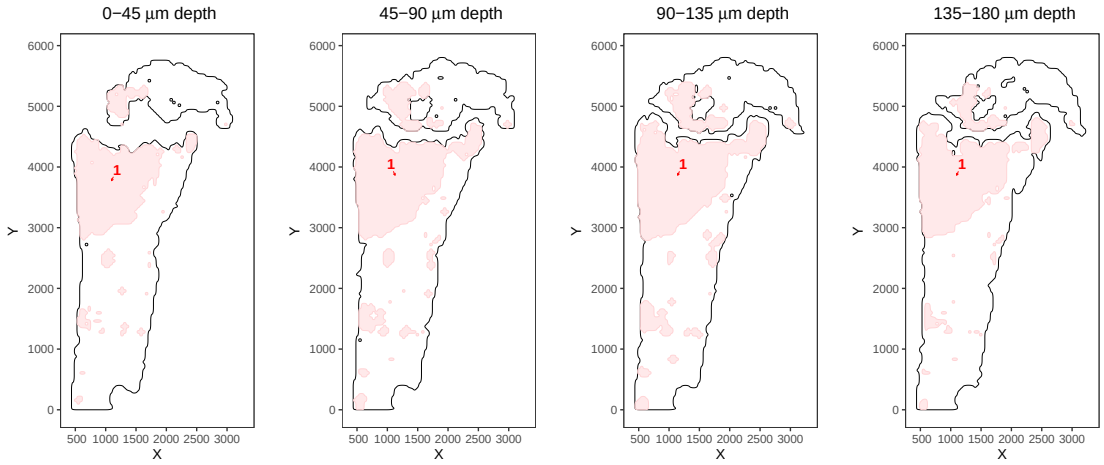
Visual inspection of Figure 2.7 showed that there were few T cells in areas with high AML cell counts. To statistically assess if the distribution of T cell counts was influenced by areas of high AML density, permutation tests were used to challenge the null hypothesis that the mean number of T cells within each cube is independent of whether the cube is within an AML cluster or in the boundary of cubes surrounding it.

The largest cluster for which the hypothesis test was rejected was marked in Figure 2.9(a). It means the mean number of T cells within and round around the Cluster 1 has statistically significant differences, where Bonferroni correction was performed. Here, the threshold for the significance of the Bonferroni-corrected p-value was $p < 0.05/18 = 0.0028$. The permutation test results were in Table 2.2. Of the remaining 17 smaller AML clusters, they fail to reject the null hypothesis. This analysis indicated that only the largest high-density AML cluster influenced the mean number of T cells.

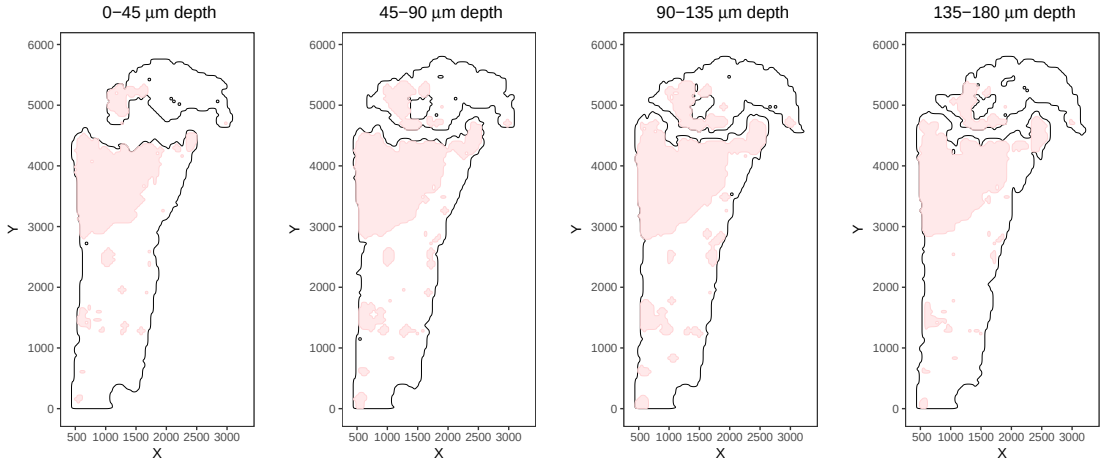
cluster number	p value for T cells	p value for MGKs
1	0.00004	0.052
2	0.28	0.76
3	0.84	0.92
4	0.13	1
5	0.13	0.44
6	0.0053	1
7	0.74	0.32
8	0.31	0.81
9	0.65	0.59
10	0.6	0.72
11	0.89	1
12	0.56	0.31
13	0.17	0.32
14	0.41	1
15	0.62	1
16	0.7	0.63
17	1	0.31
18	1	0.53

Table 2.2: P-values for spatial permutation tests.

Spatial permutation tests were also performed to test the null hypothesis that the mean number of MGKs in each cube is independent of whether the cube is within an AML cell high density area or adjacent to it. The results showed no evidence that MGKs had a different distribution within versus around any AML dense area, which has been shown



(a): Spatial permutation test between T cells and AML



(b): Spatial permutation test between MGKs and AML

Figure 2.9: Cell density dependence hypothesis test.

Two-dimensional projection visualisations that show hypothesis test results at different z-depths. The labelled pink area indicates the AML cluster where the mean number of other cells within and around the cluster is different with statistical significance. The black line represents the bone marrow boundary, and the white areas in the bone marrow are where AML cells are not clustered.

in Figure 2.9(b) and Table 2.2.

Logistic regression

To measure the likelihood of finding T cells, MGKs, or both cell types in AML cluster areas compared to non-dense AML areas, the logistic regression was used. Logistic regression is a statistical model that reports on the likelihood of an event to occur or not given a certain binary condition, in the case of the presence of T cells, MGKs or both inside versus outside AML cluster areas. From Figure 2.10, the appearance of T cells, both T cells and MGKs were affected by the AML clusters. However, in comparison, the existence of MGKs alone was not greatly influenced by AML high-density areas. The likelihood of finding cubes with T cells or both T cells and MGKs inside dense AML areas was about 80% lower than that of finding such types of cubes outside of dense AML areas and both p values $< 2 \times 10^{-16}$ were reported in Table 2.3. The likelihood of finding cubes with MGKs but not T cells inside dense AML areas was lower, but not significantly, than outside of them.

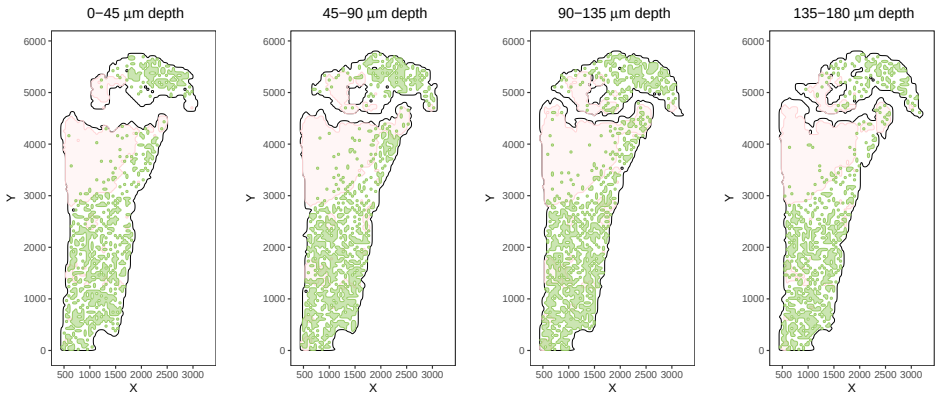
	coefficent	**Exp(coef)**	**p-value**
Presence of T cells but no MGKs	-1.62	0.198	$< 2 \times 10^{-16}$
Both T cells and MGKs	-1.63	0.196	$< 2 \times 10^{-16}$
Presence of MGKs but no T cells	-0.05	0.953	0.464

Table 2.3: Logistic regression results.

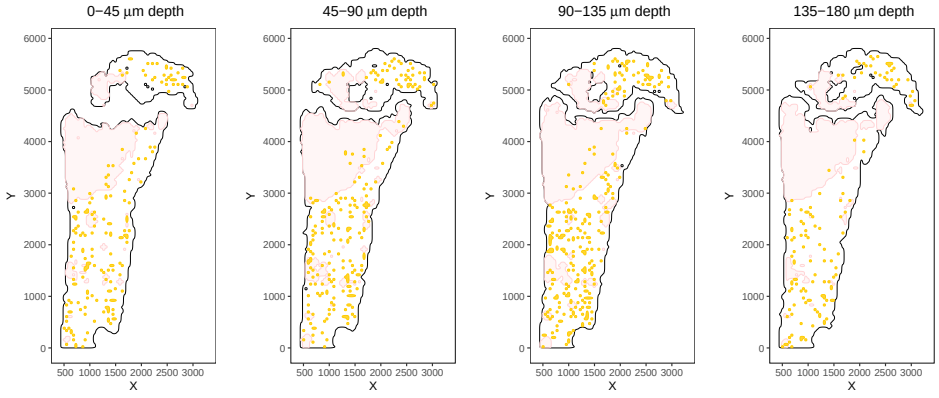
3D geographically weighted regression

In order to quantify the strength of relationships between cell counts of cells of different types, however, a statistical model is needed. To quantify the spatial relationship between T cell, MGK and AML cell counts, the 3D geographically weighted regression (GWR) model was employed. The objective of a regression model is to determine to what extent the value of response variables, such as counts of one cell type, can be explained in terms of explanatory variables, such as the cell counts of potentially related cell types, elucidating the relationship between the two. In GWR, the additional element is that the relationship can have a geographically varying dependence.

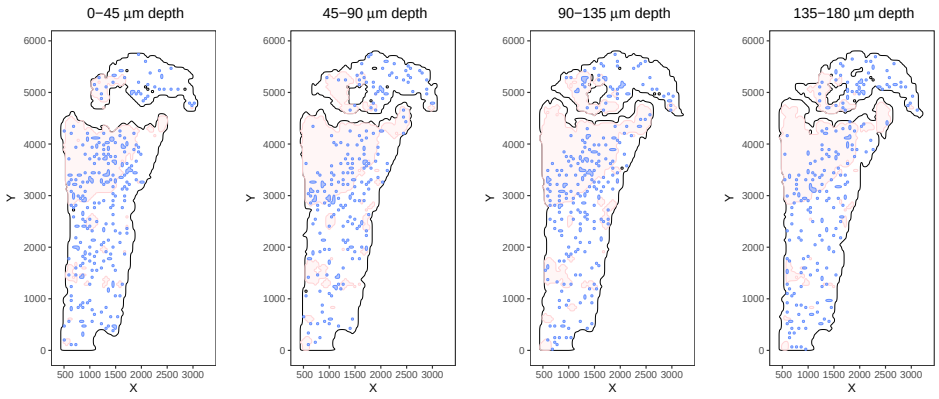
Here, the T cell count per cube was treated as the response variable, while location-dependent AML and MGK counts were used as explanatory variables. GWR models are parameterised by a spatial scale, called the optimal bandwidth, that is algorithmically determined by the data and corresponds to the maximum extent of geographic influence. For these data, that value was 455 μm, which corresponds to the width of a little over



(a): T cells and AML clusters



(b): T cells and MGKs and AML clusters



(c): MGKs and AML clusters

Figure 2.10: Logistic regression.

Two-dimensional projection visualisations that show the spatial presence of different cell types at different z-depths. The pink areas indicated the AML cluster areas, (a) the green areas were the presence of T cells but no MGKs, (b) the yellow cubes existed both T cells and MGKs, and (c) blue boxes had MGKs but no T cells.

ten cubes.

For models with no spatial component, such as linear regression, the R^2 statistic is a common measure of the quality of the model description of the data. For the spatial model GWR, the equivalent location-dependent statistic is called the local R^2. In linear regression, the coefficient of an explanatory variable is the best-fit linear multiplier that predicts the response variable given the best offset. Similarly, the GWR coefficients of location-dependent AML and MGK counts informed the multiplicative relationship between the number of T cells in a cube, and the number AML cells and MGKs contained in cubes within a $455\mu m$ range. The sign and magnitude of those coefficients captured the nature and strength of the relationship between T cell counts, and AML and MGK counts nearby. By observing how the coefficients change in a spatial context, information was extracted about geographical dependence in those relationships.

For four z-levels, Figure 2.11 (a)-(d) plotted the 3D GWR model intercept. The intercepts decreased from the distal part of the bone marrow to the growth plate region and approached zero in the high-density areas of AML, compared with the representation in Figure 2.8.

Figure 2.11 (e)-(h) indicated a positive relationship between MGKs and T cells in most areas of the bone marrow. Coefficients in the bottom areas of the bone marrow cavity where no AML clusters were negative, can be explained by the negative relationship between the number of MGKs and the number of T cells in the low AML infiltration area. The non-positive coefficients in the small upper right corner of the bone marrow cavity correspond to the location of the largest AML cluster corner.

In Figures 2.11 (i)-(l), the positive coefficient for AML effectively highlights the interaction between T cells and MGK in combating AML cells. It also revealed that in the largest AML cluster, the coefficients of MGKs and AML cells divided the cluster into two sub-regions. In the region with high AML infiltration, two patterns emerge.

One shows a weak non-positive correlation between AML cell counts and T cell counts, with T cells slightly increasing as MGKs counts rise. The other, located at the cluster's corner, demonstrated a positive correlation between AML cells and T cells, while MGKs are negatively correlated with T cells. This cellular relationship was similar to that observed in low-infiltration areas, suggesting that low or high AML cell numbers could affect the relationship between T cells and MGKs within a small range. It indicated that the 3D GWR could identify the regional variations in tissue composition.

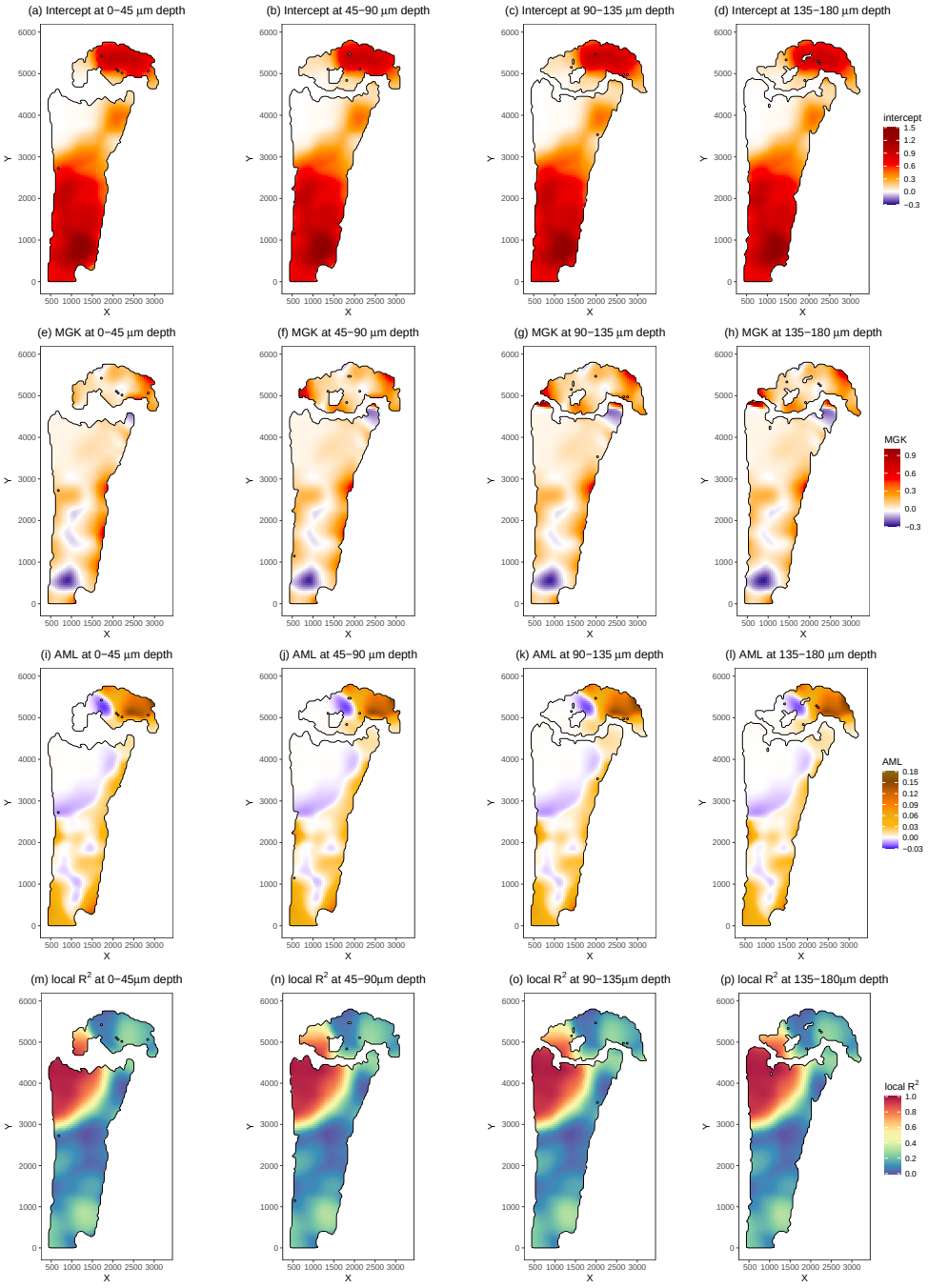


Figure 2.11: 3D statistical map of leukemic bone.

Projection visualisation for 3D spatial model results of the intercept of (a-d) the spatial model in different vertical layers. The coefficient of MGKs and the coefficient of AML cells are shown in (e-h) and (i-l), respectively, and the red part shows that the coefficient of this area is positive, the blue part is negative. The darker the colour, the larger the coefficient value. The summary statistic (m-p), local R^2, describes the model quality in each cube.

The main negative or zero positions in Figure 2.11 (i)-(l) almost coexisted with the three largest AML clusters which contained more than 1% AML cells, while the coefficients in the other small clusters were positive. This quantifies that large areas of high AML density served to exclude T cells, while smaller AML clusters do not see such a strong effect.

Figure 2.11 (m)-(p) reported the local R^2 values that provided a spatial understanding of how well GWR explains the data. Most variation occurred in the x-y plane, with little spatial variation in the z direction, which, for this bone, has the smallest extent and almost complete coverage by the large AML clusters. Matching with intuition, these showed that the model was a particularly good predictor of T cell counts in areas of the bone marrow where AML cells were densest. This is consistent with the previous results of the spatial permutation tests and logistic regression.

Taken together, the analysis of these data illustrates how methods from spatial statistics can be adapted to a 3D framework to enable the quantitative evaluation of clustering, the automatic identification of regions of high density, and the statistical assessment of dependencies between cell types.

2.5 Discussion

In this chapter, a novel and effective spatial statistical pipeline was introduced, for accessing the spatial distribution of cells at a fixed time from complex biological images. The method builds on established techniques from machine learning and spatial statistics and has a number of advantages over current approaches.

In previous studies, researchers have challenged the null hypothesis that the distribution of each individual cell population is uniform at random throughout the space through the use of a simulation method called random dots [29, 74, 137]. In a spatial analysis context, however, the null hypothesis must be conditioned to take values only within the observable, imaged, space. Although the random dots method is intuitive, there is no guarantee that the null hypothesis is physiologically viable as researchers typically cannot observe all features within the tissue with current technology, which introduces a risk of bias in the findings by increasing the likelihood of rejection. This shortcoming was circumvented in the spatial statistical pipeline, avoiding the use of simulated data, and instead addressing the question of how cells within the bone marrow are related by other approaches, including direct comparisons between multiple different cell populations.

The spatial statistics approach provides a reproducible method for quantifying the properties of individual cell types, as well as interactions between them. Simple measures evaluate the extent of spatial clustering of individual cell types, while geographically aware clustering methods then identify regions of high density.

Permutation tests provide statistical measures of high-density areas and their surroundings for assessing the relatedness of cell densities in reasons of interest, and logistic regression quantifies the extent to which the relative abundance of a cell type in high-density and non-high-density areas depends on other cell types.

Geographically weighted regression can then quantify the relationships between multiple cell types in each cube simultaneously. At its core, it is a form of regression and the output is a series of 3D maps which describe how cellular coefficients vary between different cell populations across space. It identifies areas in which spatial effects, such as those generated by an expanding malignant infiltrate, have the most influence on the presence of other cell types. It adjusts coefficients to take into account the presence of cellular interactions. Importantly, while the 3D geographically weighted regression with three cell types was demonstrated, this could be scalable to include any number of cell types.

While interest in leukaemia and T cells interactions has been growing, stemming from immunotherapy developments, MGKs are not normally studied together with the other two cell types, and the image spatial analysis highlighted some unexpected spatial relationships between the three. Consistent with previous reports of AML cells growing in localised patches within the bone marrow tissue [40, 119], several clusters of AML cells were identified. All larger clusters were adjacent to the endosteum, again consistent with previous reports that the AML model tends to grow from endosteal regions [40]. The fact that no cell type returned a negative Moran's I value is consistent with the current working hypothesis that haematopoietic cells are randomly distributed across bone marrow space, with no ordered topology.

Interestingly, while several reports describe severe loss of all haematopoietic cell types as a consequence of AML growth, our analysis highlighted that heavily infiltrated AML areas may not be homogeneous, as the bone marrow area containing the largest AML cell cluster we observed could be split into two sub-regions, one MGK type was positively correlated with T cells, while the other was negatively correlated. This observation opens the question of whether there may be major differences in the rate of loss of different

healthy haematopoietic cell types in the bone marrow as AML grows, and whether areas more greatly devoid of healthy cells represent areas where AML cells have been present for longer.

In summary, the ultimate goal of quantitative 3D imaging is to informatively summarise, in a numerical form, the vast amount of spatial and cellular information present within an image. This is particularly useful for settings where the principles regulating cell distribution are not clear, such as in bone marrow tissue.

The chapter presented here provides a framework for the analysis of complex 3D bone marrow tissue image data. The spatial statistical analytical workflow could aid in the interpretation of the spatial distribution of cells at a fixed time, which can provide a comprehensive insight into the relationships which exist within cell types in the bone marrow. It promises to uncover the principles regulating the cellular organisation of the cells responsible for the lifelong production of blood cells and their deregulation when haematological disease occurs.

3

Statistical analysis of spatio-temporal cell type distribution in 4D in vivo time-lapse data

3.1 Abstract

This chapter presents the work about intravital microscopy and spatio-temporal statistical image analysis of early engraftment of haematopoietic stem cells revealing dynamic and evolving cellular behaviours linked to progressive differentiation. This work was carried out in collaboration with partners in Professor Cristina Lo Celso's lab at Imperial College London, and the related paper is currently in preparation. My main contribution was to design and conduct a statistical framework to clean the spatio-temporal dataset, and analyse the cell type distribution by unsupervised machine learning method based on the high-dimensional data collected by our collaborators.

Haematopoietic stem cells (HSCs) have long been used in the clinic for bone marrow transplantation applications, critical for the survival of an increasingly wide range of patients with haematological, oncological and immunological pathologies [135]. Despite this, little is known about the mechanisms through which relatively few stem cells can regenerate the entire haematopoietic tissue of transplant recipients. Stem cells are located in an environment called the niche, which controls how stem cells generate, maintain, and repair tissues. To gain insights into this biological process, Prof. Lo Celso's lab used intravital microscopy of calvarium bone marrow, collecting tissue-wide images and a total of 850 hours of time-lapse data of engrafting HSCs from two days to four days following injection in lethally irradiated recipients. My contribution formed the development of methods to draw statistically meaningful conclusions from this high-dimensional spatio-temporal dataset. Analysis of the data revealed that regenerating HSCs and their immediate progeny are highly dynamic, migrating through the parenchyma at both the microscopic and near-macroscopic scales, i.e. within and in-and-out of fields of view.

Spatio-temporal statistical analysis of single cell level highlighted increasingly heterogeneous cellular behaviours over time. In particular, persistence in the vicinity of nestin-GFP perivascular cells correlated with stemness, while more heterogeneous niche interactions correlated with differentiation. The findings presented here shed light on the fundamental principles driving haematopoietic regeneration in transplantation settings.

3.2 Introduction

Haematopoietic stem cells (HSCs) play a crucial role in the adult life of vertebrates by sustaining continuous blood production throughout life. They are characterised by their multipotency and self-renewal capacities, enabling them to generate all blood cell lineages while maintaining a pool of undifferentiated stem cells [131]. This unique capability has driven the early clinical and widespread clinical application since their discovery [97, 144].

HSC transplantation is widely used as a medical treatment for various diseases, particularly haematological malignancies, as well as other blood disorders and autoimmune diseases [68]. However, despite the rapid adoption of HSC transplantation as a therapeutic intervention, many patients experience engraftment failure due to the incomplete understanding of the complex processes involved [32]. Understanding stem cell activity within the bone marrow microenvironment is essential not only for revealing the fundamental mechanisms of stem cell function but also for improving stem cell transplantation therapies and improving patient recovery [92].

Haematopoietic cell hierarchy has been extensively studied. From the top of the pyramid, dominated by HSCs, which could differentiate into multipotent progenitors (MPPs), to the bottom where are all the terminally differentiated cells [82]. Several surface markers have been identified to describe and characterise these populations phenotypically [82]. However, understanding the spatial and temporal dynamics of HSCs *in vivo* remains a considerable challenge due to the complexities of observing and analysing their behaviours within the bone marrow [13, 66].

One of the most difficult aspects of studying HSCs is their niche, the specialised microenvironment in the bone marrow, which plays an important role in HSCs maintenance, proliferation, and differentiation [89]. Interactions with components such as mesenchymal stromal cells in the niche are crucial for HSCs function [101, 108], but the mechanisms that regulate these interactions and promote HSCs survival and regenerative capacity remain unclear. Defining this is essential to learn how to maintain and expand HSCs, and how to achieve a faster bone marrow recovery following injury.

Using *in vivo* imaging is a powerful method to study the behaviour of HSCs and MPPs over time, capturing their dynamic interactions within the bone marrow niche [89]. This method allows for longitudinal observation of cells in living mice, which is extremely challenging due to the difficulty of maintaining the survival of the animals throughout the experiment. The resolution required to capture cell behaviour in the bone marrow is highly complex, as cells need to be identified by staining for specific markers, and tracking them across time and space demands specialised imaging techniques [90]. In this chapter, a combination of intravital microscopy and fluorescent staining was employed to capture the early stages of haematopoietic stem and progenitor cell (HSPC) engraftment and regeneration after lethal irradiation and transplantation.

While *in vitro* studies using biological surface markers and single-cell technology can effectively analyse cell differentiation, there remains a significant gap in studying this process directly *in vivo* [82]. The data used in this chapter were collected through cutting-edge life sciences methodologies in Prof. Lo Celso's lab. The data collection process for this work was demanding, requiring 850 hours of time-lapse imaging to observe interactions between HSPC with other cells in the niche. This resulted in the creation of a high-dimensional dataset that captured spatio-temporal variables, where the noise, missing data, or spatio-temporal inconsistency could significantly impact the analyses.

Given the challenges of acquiring this *in vivo* images and keeping mice alive while capturing detailed cellular interactions, datasets like this are valuable and require a statistical framework to accurately interpret the biological meaning. Therefore, mechanisms to automatically identify outliers so that they could be checked against original images and videos proved to be important and needed to be approached with consideration of the specific biological context to ensure data integrity.

This chapter is divided into two main sections, focusing on my contribution to this joint research endeavor: spatio-temporal data cleaning and data analysis. Both sections aim to identify key characteristics of the motion of distinct cell types.

3.3 Methods

In this section, the statistical methods used to analyse the data are described. While some of these methods may be familiar to the reader, their inclusion ensures that the thesis is self-contained and comprehensive. All programming was performed using `R` (version `R` 4.3.2).

In the dataset, there were M unique cells. Within each unit of time t_u (min), each cell's real-time coordinate position was obtained. In detail, the kth cell was recorded N_k times. At time t_j, the central position of the kth cell of type τ was $P_{k,\tau}^{t_j} = (x_{k,\tau}^{t_j}, y_{k,\tau}^{t_j}, z_{k,\tau}^{t_j})$, where $0 \leq t_j \leq N_k - 1, t_j \in \mathbb{N}$, $\mathbb{N}$ is the set of natural number, $1 \leq k \leq M, k \in \mathbb{Z}$, $\mathbb{Z}$ is the set of integers, and $\tau \in \{1, 2, 3\}$.

The distance between the kth cell of type τ at time t_{j-1} and time t_j is

$$d_{k,\tau}^{t_{j-1}t_j} = \begin{cases} \sqrt{(x_{k,\tau}^{t_{j-1}} - x_{k,\tau}^{t_j})^2 + (y_{k,\tau}^{t_{j-1}} - y_{k,\tau}^{t_j})^2 + (z_{k,\tau}^{t_{j-1}} - z_{k,\tau}^{t_j})^2} & \text{if } j \neq 0 \\ 0 & \text{if } j = 0 \end{cases}, \tag{3.1}$$

where $0 \leq t_{j-1} < t_j \leq N_k - 1$ and $t_j - t_{j-1} = t_u$.

Based on Formula 3.1, the speed for kth cell from time t_{j-1} to time t_j was $d_{k,\tau}^{t_{j-1}t_j}/t_u$ (μm/min). Let $\bar{d}_{k,\tau}$ be the average speed and $\sigma_{k,\tau}$ be the standard deviation of speed for the kth cell of type τ

$$\bar{d}_{k,\tau} = \frac{1}{N_k - 1} \sum_{j=1}^{N_k - 1} \frac{d_{k,\tau}^{t_{j-1}t_j}}{t_u},$$

$$\sigma_{k,\tau} = \sqrt{\frac{1}{N_k - 1} \sum_{j=1}^{N_k - 1} \left(\frac{d_{k,\tau}^{t_{j-1}t_j}}{t_u} - \bar{d}_{k,\tau} \right)}.$$

In order to challenge the null hypothesis that the average mean speed was independent of whether the cell type was $\tau = 1$ or the cell type was $\tau = 2$, a permutation test was used [83].

The set $A_{\bar{d}} = \left(\bar{d}_{1,1}, \bar{d}_{2,1}, \cdots, \bar{d}_{M_1,1}, \bar{d}_{1,2}, \bar{d}_{2,2}, \cdots, \bar{d}_{M_2,2}\right)$ was an ordered set, and it included $M_1 + M_2$ elements, where M_1 was the number of cells in type 1 and M_2 was the number of cells in type 2. The first M_1 elements were the mean speed records for type 1 cells, and the last M_2 elements were the mean speed records for type 2 cells. The real-valued statistic used under this null hypothesis was

$$M(A_{\bar{d}}) = \left| \frac{\sum_{k=1}^{M_1} \bar{d}_{k,1}}{M_1} - \frac{\sum_{k=1}^{M_2} \bar{d}_{k,2}}{M_2} \right|.$$

Then, create a permutation test π that reassigns labels to individual data. A reordered

observation was obtained

$$A_{\pi(\bar{d})} = \left(\bar{d}_{\pi(1,1)}, \bar{d}_{\pi(2,1)}, \cdots, \bar{d}_{\pi(M_1,1)}, \bar{d}_{\pi(1,2)}, \bar{d}_{\pi(2,2)}, \cdots, \bar{d}_{\pi(M_2,2)}\right).$$

A collection of permutation Q could be characterised so that reordering $\left\{A_{\pi(\bar{d})}\right\}_{\pi\in Q}$ was used to compare with $M(A_{\bar{d}})$. Therefore, the Monte Carlo approximation with $|Q|$ test statistics was applied to estimate the two-tailed p-value

$$p = \frac{1 + \sum_{j=1}^{|Q|} I_{|M(A_{\bar{d}})| \leq \left|M(A_{\pi(\bar{d})_j})\right|}}{|Q|},$$

where I was the indicator function.

Permutation tests and Monte Carlo approximation were also used to statistically assess the null hypothesis that the average mean speed was independent of whether the cell type was $\tau = 1$ or the cell type was $\tau = 3$, and the null hypothesis that the average mean speed was independent of whether the cell type was $\tau = 2$ or the cell type was $\tau = 3$. The threshold for significance of the Bonferroni-corrected p-value was $p < 0.05/3 = 0.0167$.

To challenge the null hypothesis that the average standard deviation of speed was independent of whether the cell type was $\tau = i$ or the cell type was $\tau = j$, where $i, j \in \{1, 2, 3\}$ and $i \neq j$. The real-valued statistic was

$$M(A_\sigma) = \left| \frac{\sum_{k=1}^{M_i} \sigma_{k,i}}{M_i} - \frac{\sum_{k=1}^{M_j} \sigma_{k,j}}{M_j} \right|.$$

Afterwards, the Monte Carlo approximation was applied to get the two-tailed p-value, and Bonferroni-corrected performed.

The confinement ratio and confinement index described the linearity of cells' movement. In detail, the confinement ratio ($c^r_{k,\tau}$) was the ratio of the distance between the first and final positions to the path length of the kth cell of type τ[64]. The confinement index ($c^i_{k,\tau}$) was calculated as the ratio of the max radius to path length for the kth cell of type τ [134]. Max radius ($\max(d)_{k,\tau}$) was the furthest distance between kth in its movements

$$\max(d)_{k,\tau} = \max\left(\sqrt{(x^{t_i}_{k,\tau} - x^{t_j}_{k,\tau})^2 + (y^{t_i}_{k,\tau} - y^{t_j}_{k,\tau})^2 + (z^{t_i}_{k,\tau} - z^{t_j}_{k,\tau})^2}\right),\ 0 \leq t_i, t_j \leq N_k - 1,$$

and path length ($\sum(d)_{k,\tau}$) was the sum of the movement of the kth cell of type τ

$$\sum(d)_{k,\tau} = \sum_{t_j=1}^{N_k-1} d_{k,\tau}^{t_{j-1}t_j} .$$

The confinement ratio were defined as

$$c_{k,\tau}^{r} = \frac{d_{k,\tau}^{t_0 t_{N_k-1}}}{\sum(d)_{k,\tau}} ,$$

and confinement index was

$$c_{k,\tau}^{i} = \frac{\max(d)_{k,\tau}}{\sum(d)_{k,\tau}} .$$

In order to quantify the relationship between the max radius and the distance between the first and final position, the Pearson correlation coefficient was applied, and the range of the Pearson correlation coefficient was $[-1, 1]$ [128]. It was performed by `'cor'` function with method `'pearson'` in `'R'`. The closer the coefficient is to 1, the more correlated the two variables are.

To challenge the null hypothesis that the average value of confinement ratio was independent of the cell types, and permutation test were used. The real statistic for the $\tau = i$th cell type and $\tau = j$th cell type, where $i, j \in \{1, 2, 3\}$ and $i \neq j$, was

$$M(A_{c^r}) = \left| \frac{\sum_{k=1}^{M_i} c_{k,i}^{r}}{M_i} - \frac{\sum_{k=1}^{M_j} c_{k,j}^{r}}{M_j} \right| .$$

Monte Carlo approximation was employed to get the two-tailed p-value, and Bonferroni correction was used to change the threshold for significance.

Permutation tests and Bonferroni correction were also used to statistically assess whether the average value of confinement index was independent of the cell types. The two-tailed p-values were calculated by Monte Carlo approximation with 250,000 sampling.

The angle could also be used to analyse the linearity of cell movement [134]. In detail, the vector $\overrightarrow{P_{k,\tau}^{t_{j-1}} P_{k,\tau}^{t_j}}$ means the movement for the kth cell from time t_{j-1} to time t_j. Hence,

the angle for the kth cell of type τ at time t_j was defined by

$$cos(\theta_{k,\tau}^{t_j}) = \frac{\overrightarrow{P_{k,\tau}^{t_{j-1}} P_{k,\tau}^{t_j}} \cdot \overrightarrow{P_{k,\tau}^{t_j} P_{k,\tau}^{t_{j+1}}}}{\left|\overrightarrow{P_{k,\tau}^{t_{j-1}} P_{k,\tau}^{t_j}}\right| \left|\overrightarrow{P_{k,\tau}^{t_j} P_{k,\tau}^{t_{j+1}}}\right|},$$

and

$$sin(\theta_{k,\tau}^{t_j}) = \frac{\overrightarrow{P_{k,\tau}^{t_{j-1}} P_{k,\tau}^{t_j}} \times \overrightarrow{P_{k,\tau}^{t_j} P_{k,\tau}^{t_{j+1}}}}{\left|\overrightarrow{P_{k,\tau}^{t_{j-1}} P_{k,\tau}^{t_j}}\right| \left|\overrightarrow{P_{k,\tau}^{t_j} P_{k,\tau}^{t_{j+1}}}\right|},$$

if $\left|\overrightarrow{P_{k,\tau}^{t_{j-1}} P_{k,\tau}^{t_j}}\right| = 0$ or $\left|\overrightarrow{P_{k,\tau}^{t_j} P_{k,\tau}^{t_{j+1}}}\right| = 0$, $\theta_{k,\tau}^{t_j} = 0$.

The arrest coefficient ($a_{k,\tau}$) was the percentage of time for the instantaneous velocity of kth cell of type τ was less than $2\mu m/\text{min}$ [69]

$$a_{k,\tau} = \frac{\sum_{j=1}^{N_{k,\tau}-1} I_{k,\tau}^{t_{j-1}t_j}}{N_k - 1} \times 100\%,$$

where $I_{k,\tau}^{t_{j-1}t_j} = \begin{cases} 1, \frac{d_{k,\tau}^{t_{j-1}t_j}}{t_u} \leq 2 \\ 0, \frac{d_{k,\tau}^{t_{j-1}t_j}}{t_u} > 2 \end{cases}$ is an indicator function.

Permutation tests and Monte Carlo approximation were used to statistically challenge the difference in arrest coefficient between different cell types. The null hypothesis was the mean of the arrest coefficient was independent of the cell type. Bonferroni correction was performed.

Inspired by the arrest coefficient, the nestin arrest coefficient (NAC) was defined as the proportion of time a cell was less than 25 μm away from the nestin-GFP+ cells. The shortest distance from the kth cell of type τ to the nestin-GFP+ cells at each time was recorded as $\left\{d_{n_k,\tau}^{t_0t_1}, d_{n_k,\tau}^{t_1t_2}, \cdots, d_{n_k,\tau}^{t_{N_k-1}t_{N_k}}\right\}$. Hence, the NAC ($n_{k,\tau}$) for the kth cell was

$$n_{k,\tau} = \frac{\sum_{j=1}^{N_k-1} I_{n_k,\tau}^{t_{j-1}t_j}}{N_k - 1} \times 100\%,$$

where,

$$I_{n_k,\tau}^{t_{j-1}t_j} = \begin{cases} 1, \ d_{n_k,\tau}^{t_{j-1}t_j} \leq 25 \\ 0, \ d_{n_k,\tau}^{t_{j-1}t_j} \geq 25 \end{cases}.$$

Permutation tests were used to detect changes in NAC between different cell types. The null hypothesis for the test is that the mean value of the NAC is independent of the cell types. The Monte Carlo approximation with 250,000 test statistics was employed to calculate the two-tailed p-values with a Bonferroni-corrected p-valued threshold of $p < 0.05/3 = 0.0167$.

To compare the cell movement behaviours close to the nestin-GFP+ cells and far away from the nestin-GFP+ cells, the permutation tests were used. For the cell movement behaviour like arrest coefficient, the null hypothesis that the mean value of the arrest coefficient for cells in type $\tau = i$ was independent of whether they located close to the nestin-GFP+ cells or not.

The ordered set $A_{a_{\tau=i}} = \left(a_{n_1,i}, a_{n_2,i}, \cdots, a_{M_{n_i},i}, a_{1,i}, a_{2,i}, \cdots, a_{(M_i - M_{n_i}),i}\right)$, and it included M_i elements. The first M_{n_i} elements were the arrest coefficient for cells close to the nestin-GFP+ cells, and the last $M_i - M_{n_i}$ element includes all arrest coefficient for cells far away from the nestin-GFP+ cells. The real-valued statistic under this null hypothesis was

$$M(A_{a_{\tau=i}}) = \left| \frac{\sum_{k=n_1}^{M_{n_i}} a_{k,i}}{M_{n_i}} - \frac{\sum_{k=1}^{(M_i - M_{n_i})} a_{k,i}}{(M_i - M_{n_i})} \right| .$$

A permutation π was created, and it reassigns labels to individual vectors. A recorded observation set is obtained $A_{\pi(a_{\tau=i})} = \Big(a_{\pi(n_1,i)}, a_{\pi(n_2,i)}, \cdots, a_{\pi(M_{n_i},i)}, a_{\pi(1,i)}, a_{\pi(2,i)}, \cdots, a_{\pi((M_i - M_{n_i}),i)}\Big)$. Then, the empirical distribution of a collection of permutations was used to compare with $M(A_{a_{\tau=i}})$. Therefore, the Monto-Carlo approximation was applied to calculate the two-tailed p-value.

All displacements for kth cell is recorded as $\left\{d_k^{t_0 t_1}, d_k^{t_1 t_2}, \cdots, d_k^{t_{N_k - 1} t_{N_k}}\right\}$. The mean of all displacements was $(\sum_{j=1}^{N_k - 1} d_k^{t_{j-1} t_j})/(N_k - 1)$. The autocovariance function at lag η, where $0 \leq \eta \leq N_k - 1$, $\eta \in \mathbb{Z}$, was

$$f_k(\eta) = \frac{\sum_{j=\eta+1}^{N_k - 1} \left(d_k^{t_{j-1} t_j} - \frac{\sum_{j=1}^{N_k - 1} d_k^{t_{j-1} t_j}}{N_k - 1}\right) \left(d_k^{t_{j-1-\eta} t_{j-\eta}} - \frac{\sum_{j=1}^{N_k - 1} d_k^{t_{j-1} t_j}}{N_k - 1}\right)}{N_k - 1} .$$

The autocorrelation function (ACF) for the kth cell at lag η is $f_k(\eta)/f_k(0)$. Furthermore, the average autocorrelation function (average ACF) was a statistic used to describe all

cells' movement memory at lag $\dot{\eta}$ was defined as

$$\bar{f}(\dot{\eta}) = \frac{\sum_{i=1}^{M} \widehat{f}_k(\dot{\eta})}{\sum_{i=1}^{M} I_k^f(\dot{\eta})},$$

where $\dot{\eta} \in \mathbb{Z}$, $\widehat{f}_k(\dot{\eta}) = \begin{cases} 0, \dot{\eta} > N_k - 1 \\ f_k(\dot{\eta})/f_k(0), \dot{\eta} < N_k \end{cases}$, and $I_k^f(\dot{\eta}) = \begin{cases} 0, \dot{\eta} > N_k - 1 \\ 1, \ \dot{\eta} < N_k \end{cases}$.

Clustering is an unsupervised learning technique that can divide data into clusters based on their similarity [5, 7]. To explore cell behaviour patterns from spatio-temporal data, Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP), a dimension reduction method based on manifold theory and topological data analysis, was used [99]. UMAP showed better performance in both local and global structures than other clustering methods [11]. These clustering and visualisation results were performed in '`R`' by '`umap`' package (version 0.2.10.0).

3.4 Spatio-temporal data cleaning

In this section, the biological experimental background is introduced, and, based on it, the cleaning procedures we created for the spatio-temporal high-dimensional dataset are performed.

3.4.1 Biological background

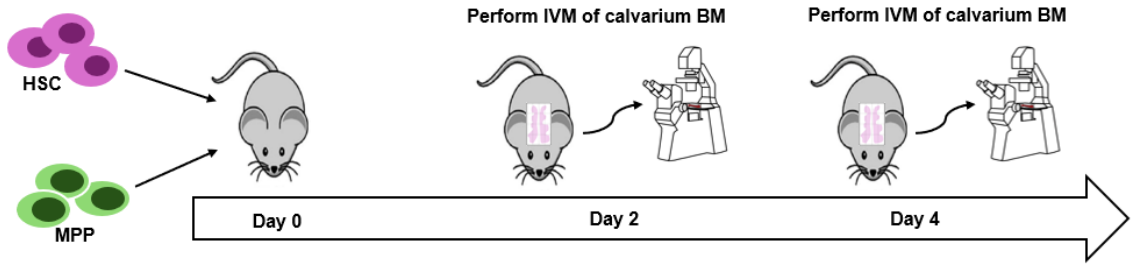


Figure 3.1: Haematopoietic stem and progenitor cells *in vivo* experiment sketch.

All biological experiments were designed and conducted by Prof. Lo Celso's lab, with a high-level sketch shown in Figure 3.1. Our collaborators injected fluorescently stained cells, haematopoietic stem cells (HSCs: Lin-, c-Kit+, Sca-1+, CD48-, CD150+, Flk2-, CD34-) and multipotent progenitors (MPPs: Lin-, Ckit+, Sca1+, CD48+, CD150-, CD34+), into the lethally irradiated mice. Recipient mice were either wild type, nestin-GFP or Flk1-GFP transgenic mice, the latter two allowing simultaneous observation

of injected cells and either nestin-GFP positive periarteriolar (nestin-GFP+ cells) or endothelial cells (ECs). Then, intravital microscopy (IVM) of the calvarium bone marrow was performed, and time-lapse data videos were collected from the recipients on the second and fourth days after the injection.

Our collaborators at Imperial College London performed the intravital microscopy to image the calvarium bone marrow *in vivo*. They imaged an area of approximately 200 $\times$ 200 $\times$ 36 μm^3 every three minutes, with one single tracking time-lapse video extending up to 684 minutes (over 11 hours). The total tracking time was about 850 hours. The spatio-temporal imaging data were then analysed with '`Imaris`' software to collect the three-dimensional coordinates of each cell at each time point. Based on the time of data collection and fluorescence staining, the dataset includes three types of cells, two days after HSCs transplantation, two days after MPPs transplantation and four days after HSCs transplantation. In the following analysis, these cell types are referred to as HSCDay2, MPP, and HSCDay4, respectively.

3.4.2 Dataset description

In the original dataset provided by collaborators, there were 17302 records and 23 variables, and each variable has been explained in Table 3.1. In this subsection, the format check and completeness check of the dataset were performed to improve the quality of the data [55].

Each unique cell in the dataset was identified by variables *Sort*, *Position*, *TrackID*, *Day*, *CellType*, *Time*, and *Time.original*. The variable *Time* recorded the start time of the video, while *Time.original* indicated when each cell first appeared in the field of view. The variable *Time.orginal* effectively captured cells that entered the view midway through the recording. Based on these variables, the original dataset contained 282 cells.

The variables *nestin* and *Ncontact* were used to describe the relationship between each cell and the nestin-GFP+ cells in the view. The Euclidean distance from the cell to the nearest boundary of a nestin-GFP+ cell was recorded in *nestin*. In 0.2% of the records, the distance was recorded as '> 140' and in 3% of them as '> 30', and they were manually added by our collaborators. Depending on whether the cells were in contact with nestin-GFP+ cells, collaborators also manually added the variable *Ncontact*. This is a binary variable, where '1' indicates that the cell is in contact with a nestin-GFP+ cell, '0' indicates no contact, and 'NA' means that nestin-GFP+ cells are not seen in the field of view.

Variable	Detail	Description
Sort	Integer	Positioning the field of view for each single cell
position	Integer	Positioning the field of view for each single cell
TrackID	Integer	Identifying each single cell in the same field of view
Day	Integer	day2, day4
Celltype	character	HSC, MPP
Time	Integer	0-288
Time.orginal	Integer	1-229
nestin	Numerical variable	26% NA, distance from cell to nestin-GFP+ cells
Ncontact	Dummy variable	26% NA, contact between cell and nestin-GFP+ cells
endothelium	Numerical variable	95% NA
Econtact	Dummy variable	95% NA
osteoblast	Numerical variable	99% NA
bone	Numerical variable	99% NA
X	Numerical variable	Three-dimensional coordinates
Y	Numerical variable	Three-dimensional coordinates
Z	Numerical variable	Three-dimensional coordinates
WvWv	Dummy variable	Biological information
Nestin	Dummy variable	Biological information
Col2.3	Dummy variable	Biological information
Flk1	Dummy variable	Biological information
manualTrack		Biological comments
Comments		Biological comments
X.1		Biological comments

Table 3.1: Variables in the original dataset.

Of cells injected into nestin-GFP+ reporter mice, 213 cells showed fields of view with nestin-GFP+ cells while 14 cells did not. In addition, a total of 55 cells were tracked in non-nestin-GFP+ backgrounds.

Similarly, variables *endothelium* and *Econtact* reported the distance between endothelium cells and cells, and whether the HSCs or MPPs touched with endothelium cells. The variables *bone* and *osteoblast* indicated the distance between bone and the cells. However, these variables had over 95% missing data. Due to the paucity of data, this information was not used in subsequent analyses.

There were some obvious recording errors in the original dataset, leading to data formatting issues, such as question marks appearing in numerical variables. There were 21 records with data format errors and 16 records with obvious human recording errors.

3.4.3 Dataset visualisation

After checking the format and completeness of the dataset, the visualisation of spatio-temporal data could effectively identify the noise, which includes errors or outliers [55]. The error data corresponds to records that are impossible, while outliers are statistically unusual and therefore require further investigation. This part of the data cleaning is divided into spatial visualisation and spatio-temporal visualisation of cells' behaviour to find error data or outliers [20].

Spatial data cleaning

The global data cleaning was based on each cell's displacement over a 3-minute period, which is shown in Figure 3.2. The maximum displacement observed within three minutes was 784.52 μm, which exceeded the entire observable field of view ($200 \times 200 \times 36 \mu m^3$). This was an obvious data recording error. Based on this observation, several error data and outliers also were identified, and my collaborators and I manually inspected and corrected them one by one.

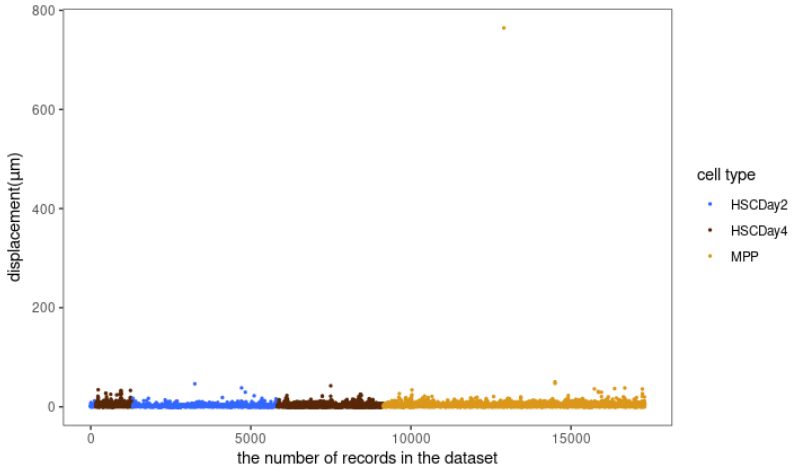


Figure 3.2: Cell displacements within three minutes in the original dataset.

Displacements of all 282 cells within three minutes, different colours represent different cell types.

Researchers have suggested that as injection time increases, HSCs and MPPs undergo differentiation into different cell types, leading to disparate displacement behaviours at different time points [46]. To account for this variability, all cells were manually examined and categorised into three groups: HSCDay2, MPP, and HSCDay4. The displacements for each cell type were then sorted from high to low and rechecked using the software

ImageJ. Figure 3.3 illustrated the displacement of each cell at each three minutes after this manual verification. Compared with Figure 3.2, this figure shows the result after spatial data cleaning, and the noise has been removed. The maximum displacement of the cell in three minutes was about 25 μm.

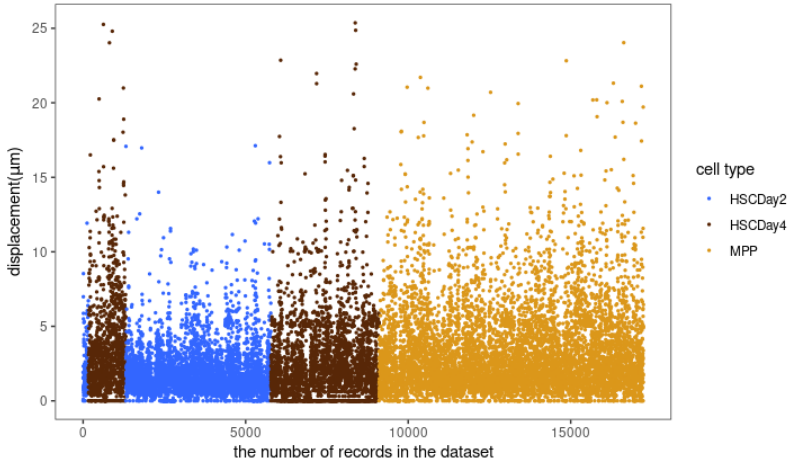


Figure 3.3: Cell displacements within three minutes in the cleaned dataset.

Displacements of all 282 cells during three-minute intervals after manual double-check, different colours represent different cell types.

Spatio-temporal data cleaning

For cells with *nestin* information, the variance in distance to nestin-GFP+ cells was analysed to explore the relationship between nestin-GFP+ cells and HSPC. A higher variance in distance indicated larger movement changes between the nestin-GFP+ cells and the cell of interest. Given that nestin-GFP+ cells exhibited minimal movement in the video and the displacement of the cell of interest every three minutes has been shown in Figure 3.3. High variance data suggested that the *nestin* information may be inaccurate, and this observation highlighted the need for further verification, helping to identify cells with potentially inconsistent *nestin* information.

Figure 3.4 showed the variance in distance to nestin-GFP+ cells for each cell in the original dataset. Based on this analysis, 20 cells were removed because their *nestin* distances were recorded as '> 30' or '> 140'. During the rechecking process, a dummy variable 'N25' was introduced for ease of recording and rechecking. This variable was set to '0' if the distance to nestin-GFP+ cells was less than or equal to 25 μm, and '1' if the distance exceeded 25 μm. Additionally, the *nestin* information for 13 cells was corrected. The reason for 25μm is explained in Section 3.5.2.

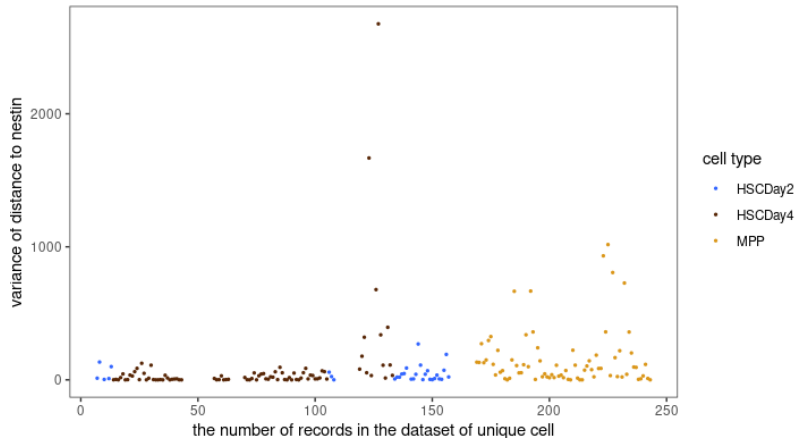


Figure 3.4: Variance of distance to nestin for each cell.

Each point represented the change in distance between the cell and nestin-GFP+ cells during movement. Twenty cells that had *nestin* information like were '> 30' or '> 140' were removed. Different colours represent different cell types.

After data cleaning, the final dataset had 17237 records, 24 variables and 282 unique cells. Of these, 12330 records (213 cells) had *nestin* information, and 4907 records (69 cells) did not have *nestin* information.

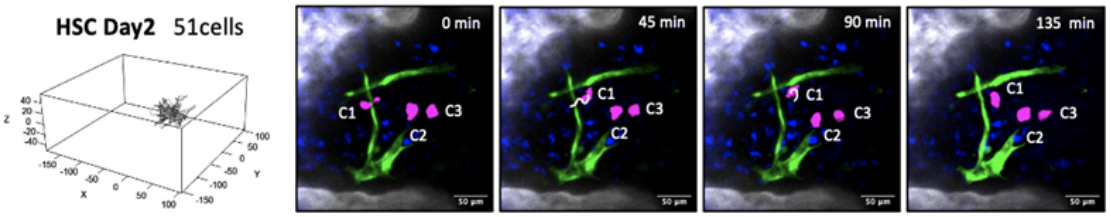
3.5 Spatio-temporal data exploratory analysis

In this section, the cleaned high-dimensional data are used for exploratory statistical data analyses. We created and executed data transformation procedures to characterise the behaviour of different types of cell distributions after different transplant times.

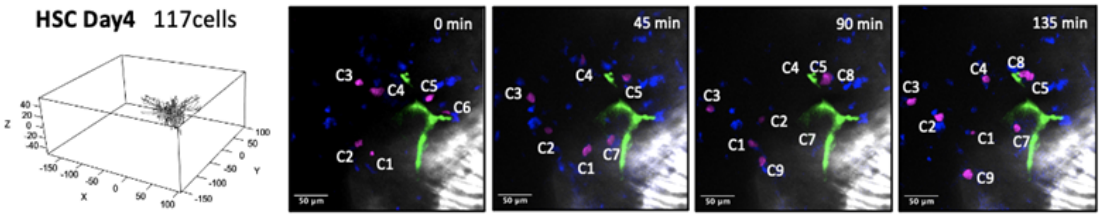
3.5.1 Time-lapse of Day2 and Day4 transplanted HSCs reveals significant dynamics and heterogeneity

Representative examples of fields of view of each cell type at selected time points are shown in Figure 3.5(a) (HSCDay2), Figure 3.5(b) (HSCDay4) and Figure 3.5(c) (MPP). Comparing the spatial and temporal trajectory of cells, HSCDay2 had the smallest cell movement range, HSCDay4 displayed a mix of behaviours, with some cells staying most of the time localising in the same point but still wiggling and producing protrusions, while others migrated along the field of view, and MPP exhibited the largest movement range.

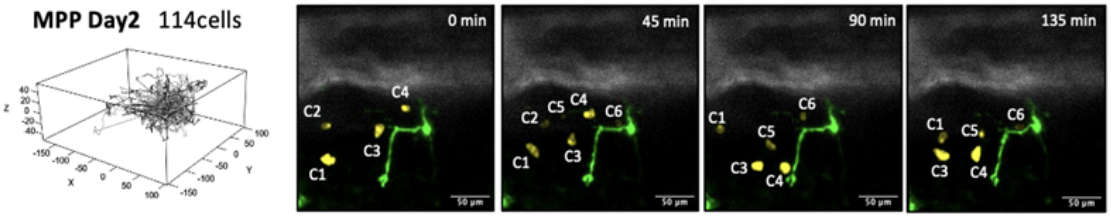
The phenomenon was also found in the screenshots in Figure 3.5. Three HSCDay2 cells moved in a small range. For HSCDay4, six cells were observed at the beginning, and some



(a): HSCDay2 movements in the example field of view



(b): HSCDay4 movements in the example field of view



(c): MPP movements in the example field of view

Figure 3.5: The three-dimensional cell trajectory obtained by time-lapse photography and the screenshots at the 0 minute, 135 minute, 270 minute and 405 minute of the time-lapse photography.

All cells were relocated to start at the origin point (0, 0, 0) for consistency in visualisation. The three-dimensional images of the three cell types share the same image range. In the screenshots, the red fluorescent cells represent HSC before transplantation, the yellow cells represent MPP, and the green cells represent nestin-GFP+ cells. The black background is the bone marrow.

cells moved in a small range like HSCDay2. Instead, some HSCDay4 cells displayed large movement trajectories, with cells constantly entering the field of view and cells jumping out of the field of view, like MPP cells.

All the tracks were shown in Figure 3.6 starting at the specific time point when a cell appears in a field of view. When tracking MPP cells, many more cells appeared at different time points during the acquisition, highlighting their greater mobility in comparison to HSCDay2 and HSCDay4.

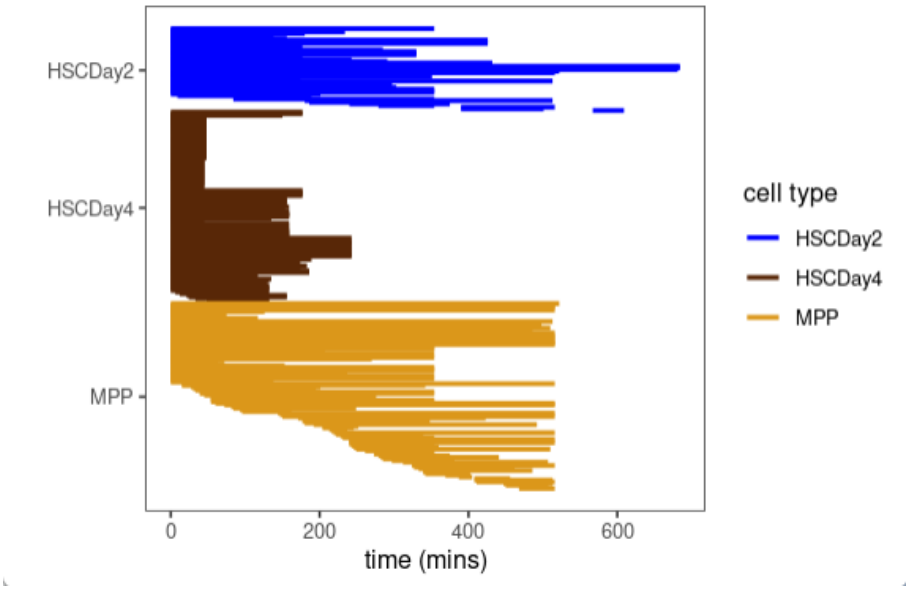


Figure 3.6: Track swimmer plot.

The length of each line represents how long each cell was tracked. The start of the line is from the specific time point when the cell appeared in the field of view. Different colours represent different cell types.

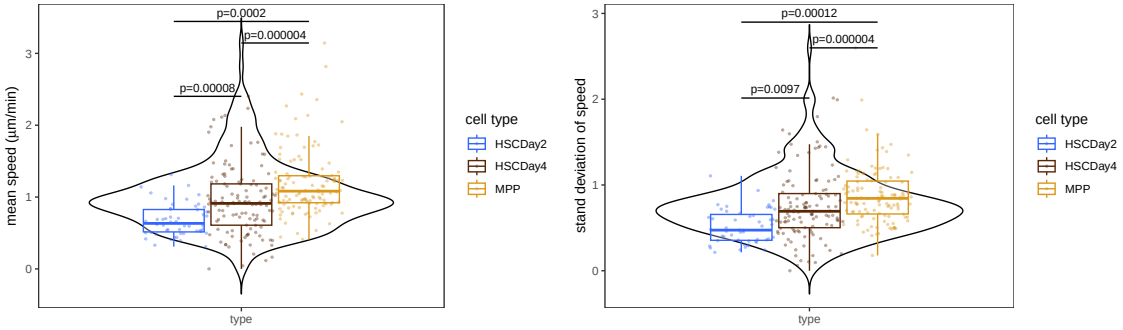
3.5.2 Analysis of tracks reveals an evolving behaviour from Day2 to Day4 and HSCs to MPPs

To identify potential specific cell behaviours associated with differentiation status, several statistical variables were analysed and compared between the three populations of interest HSCDay2, HSCDay4 and MPP. Mathematical definitions of variables are provided in Section 3.3.

HSCs two days after transplantation exhibited the lowest mean speed, consistent with

previously published data [46]. An increase in mean speed was observed in HSCDay4, and MPP cells showed the highest mean speed values in Figure 3.7(a). In order to challenge the null hypothesis that the average of mean speed was independent of cell type, permutation tests and the Bonferroni correction were performed. The threshold for significance of the Bonferroni-corrected two-tailed p-value was $0.05/3 < 0.0167$. A Monte Carlo approximation with 250,000 test statistics was used to estimate the two-tailed p-value which has been reported in Figure 3.7(a). All null hypotheses were rejected, meaning that the mean speed was significantly different between different cell types. The boxplots in the figure revealed that HSCDay4 mean speed has the largest interquartile range, indicating that the average speed of HSCDay4 has a relatively large range of variation.

Figure 3.7(b) also showed that HSCDay4 displayed a higher standard deviation, reflecting greater heterogeneity in cell movement. In contrast, HSCDay2 had a lower standard deviation than the other groups, suggesting more consistent movement. The p-values from the permutation tests and Monte Carlo approximation confirmed that the standard deviations across different cell types were significantly different. The results of the permutation tests illustrate the heterogeneity of different cell type movements.



(a): mean speed (b): standard deviation of speed

Figure 3.7: Mean and standard deviation of cell's speed.

The black violin plots show the distribution of mean speed (a) and standard deviation of speed (b) for all cells, and different colour boxplots show the statistical summary for different cell types. Each point represents the movement behaviour of each cell.

To provide an understanding of whether the movement of each individual cell was moving or pausing, the arrest coefficient was calculated as the proportion of time during which the instantaneous speed of the cell was less than 2 μm/min [69]. Figure 3.8 showed that

HSCDay2 had the highest arrest coefficient, which meant that most HSCDay2 cells moved slowly or paused. HSCDay4 had a lower median arrest coefficient than HSCDay2, and most MPP had arrest coefficients less than 100%, indicating that HSCDay2 spent more time at the same location than HSCDay4 and MPP.

Permutation tests were also used to challenge whether there were differences in the arrest coefficient across different cell types. The null hypothesis was the mean arrest coefficient was independent of cell types. Monte Carlo approximation with 250,000 test statistics was applied to estimate the two-tailed p-value, as reported in Figure 3.8. The threshold for significance of the Bonferroni-corrected p-value was $p < 0.05/3 = 0.0167$. The test results indicated that the arrest coefficient was significantly different between different cell types. HSCDay2 exhibited more limited movement compared to HSCDay4, and HSCDay4 moved more slowly than MPP. These findings suggested that the movement behaviour of HSCDay4 was more closely aligned with that of MPPs than that of HSCDay2 cells.

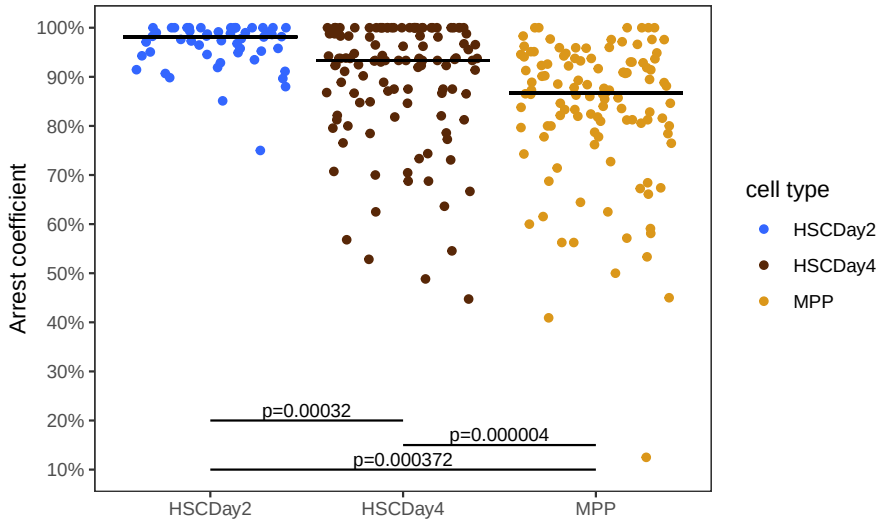


Figure 3.8: Arrest coefficient of different cell types.

Each point represents the arrest coefficient of a single cell, and the different colours show different cell types. The black line is the median arrest coefficient for different cell types.

Two variables, the confinement ratio and the confinement index, were used to describe the linearity of cell movement [64, 134]. The range of the two variables is 0 to 1, with values closer to 1 indicating more linear movement. The confinement ratio was defined

as the ratio of the distance between each cell's first position and the final position to the path length. Figure 3.9(a) showed the trajectory of a cell, where the path length was the sum of motion trajectory for this cell. The l_r, the red line in Figure 3.9(b), was the distance from the cell's first position to the final position. The confinement index is the ratio of maximum radius to path length, and the red line l_i in Figure 3.9(c) represented the max radius for the cell movement.

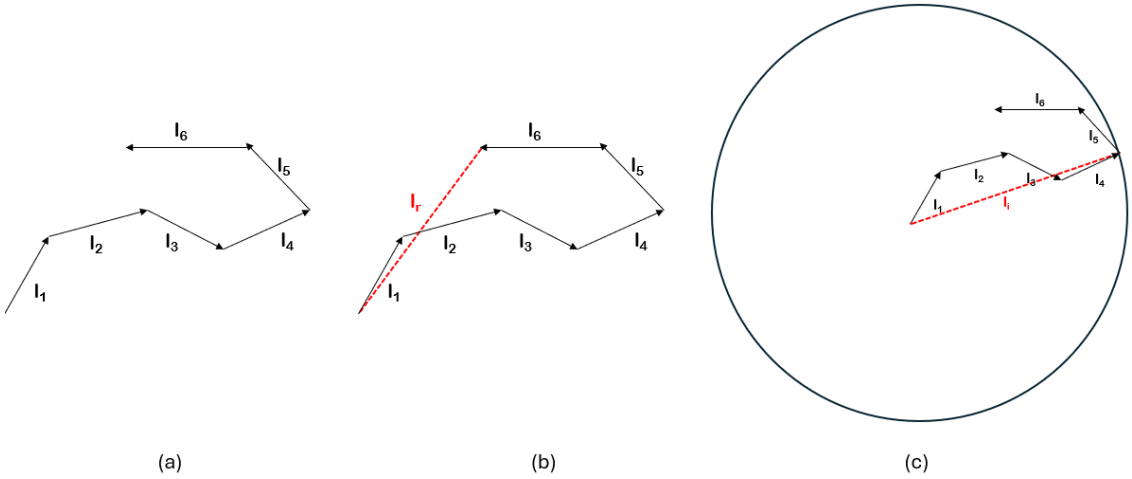
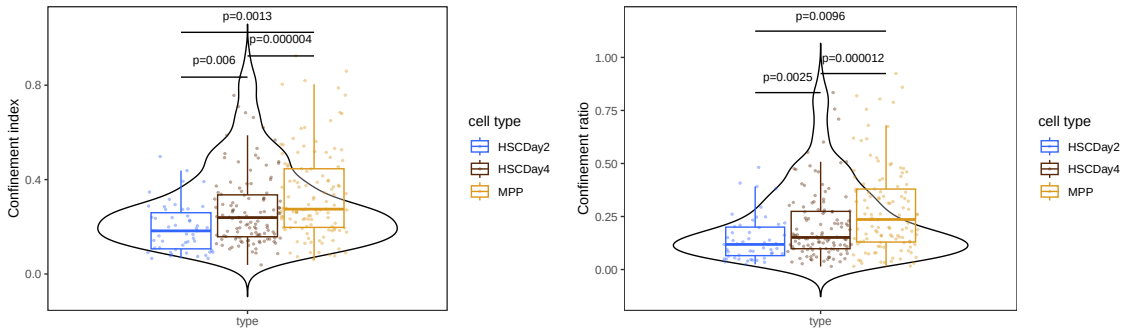


Figure 3.9: Trajectory of cell.

Path length of a cell was shown in (a). The distance from the initial position to the final position (b) and the max radius of the circle (c) were in red lines.

Figure 3.10 showed that both the confinement index and confinement ratio were larger for MPP cells than the HSCs. It displayed that as cells differentiate, the MPP cells movement became more linear. Permutation tests also suggested that the confinement index and confinement ratio were statistically independent of cell type at the Bonferroni-corrected significance level $p < 0.05/3 = 0.0167$. The permutation test results of both the confinement index and confinement ratio showed heterogeneity of linearity movement among different cell types. Since no obvious difference could be seen in Figure 3.10(a) and Figure 3.10(b), these two variables were further compared in detail.

Figure 3.11 compared the difference between the maximum radius of the cell trajectory and the distance from the first position to the final position. The colour of the point in the figure indicated the length of the time that the cell was tracked. It demonstrated for each cell, the maximum radius was equal to or larger than the distance between the first and final positions. The Pearson correlation coefficient between these two variables is 0.938, which also reveals a strong similarity between the confinement index and confinement



(a): Confinement index (b): Confinement ratio

Figure 3.10: Linearity of cell: confinement index and confinement ratio.

The black violin plots show the distribution of confinement index (a) and confinement ratio (b) for all cells, and different colour boxplots show the statistical summary for different cell types. Each point represents the information of each cell.

ratio [151].

Compared with the confinement ratio, the confinement index effectively captured the range of the cell's movement during the recording time. In detail, the confinement ratio might fail to distinguish between cases where a cell moved extensively but finally returned to its original position.

Other research groups have previously reported using cell movement angles to analyse the linearity of cell movement [134]. Based on the data, cell movement was divided into 360 degrees, and the vector method was used to calculate the angle of each movement. A rose plot was then utilised to visually represent the angular trends of cell movement. To create the rose plot, first, a histogram of angle data was generated by dividing the circular range into 360 bins, and then the frequency of data in each bin was plotted as bars radiating outward from a central point, ensuring that the start and end points of the x-axis were connected to form a continuous circular plot.

Figure 3.12 displayed the angles of all cell movements in different cell types. When cells were pausing, the movement angles were 0. This was also the reason why there were the more movements at 0 degrees. Beyond 0 degrees, it was not found that cells tend to move in a certain direction. The reason was due to spatial location extraction. In the process of obtaining the coordinates of cells, since cells had irregular shapes, the position

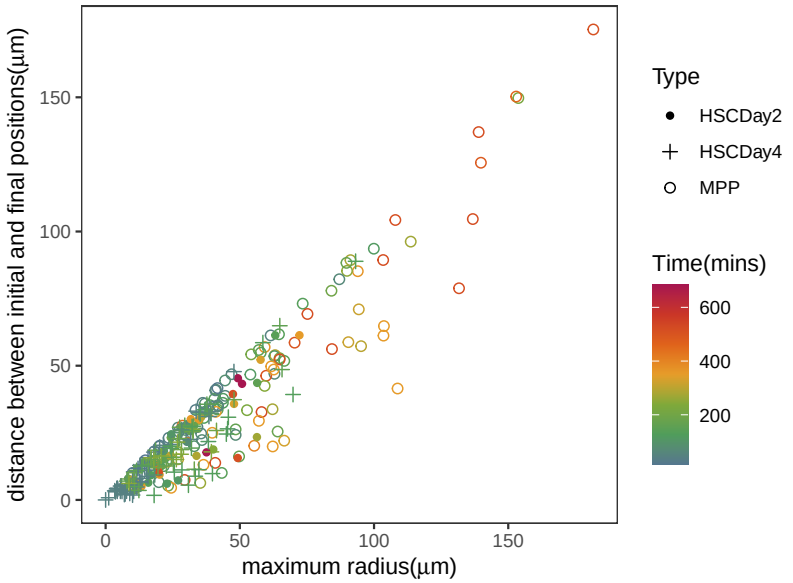


Figure 3.11: Compare max radius and distance between the first to final positions.

The x-axis is the maximum radius for each cell and the y-axis is the distance between the initial cell positions and the final cell positions. The different shapes of the cell show the cell type, and the colours show the length of time that the cell was tracked.

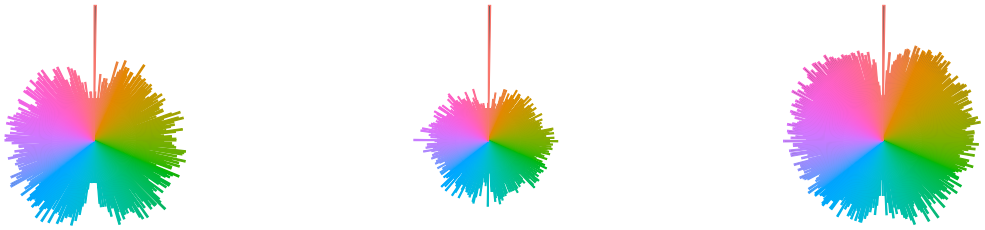
of any point on the cell (with a diameter of around 10 μm) was recorded as the real-time coordinate. The sensitivity of angles to coordinate noise may lead to errors, causing some findings to be missed. Therefore, the confinement index was most suitable for describing the linearity of cell movement in this research.

Figure 3.13 illustrated the average autocorrelation function for different cell types, and it suggested there was no temporal correlation for different cell types. This also meant that the three cell types did not exhibit memory in their movement patterns within the three-minute intervals observed.

Through the analysis of cell behaviours in this section, different types of cells had significant differences in their movements and cell movements have no memory. Next, the relationship between nestin-GFP+ cells and HSPC will be examined.

3.5.3 Analysis of transplanted HSPCs interactions with nestin-GFP+ cells

To understand whether these differential dynamics could be regulated by the bone marrow niche, HSPC (HSCs and MPPs) were transplanted into nestin-GFP transgenic reporter mice by our collaborators at day 0. Nestin cells are a subpopulation of mesenchymal stem



(a): HSCDay2 angle (b): HSCDay4 angle (c): MPP angle

Figure 3.12: Linearity of cell: rose plot analysis.

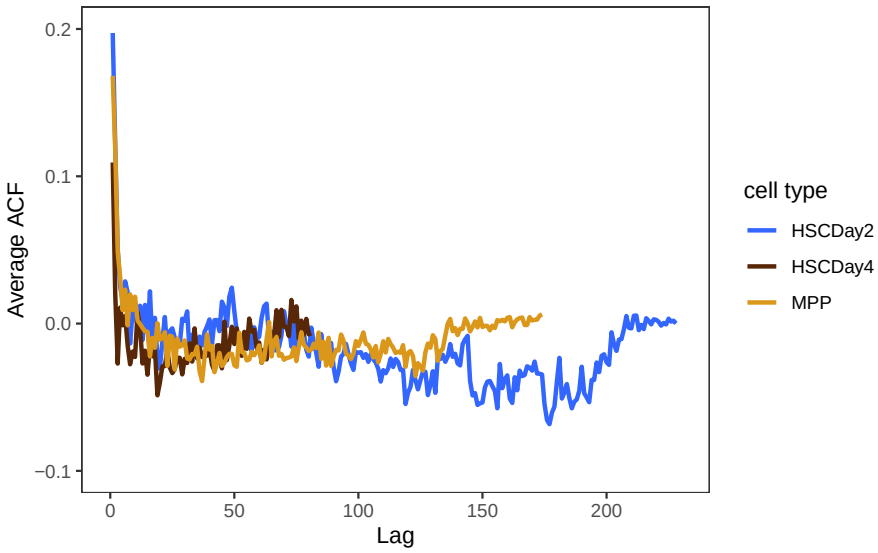


Figure 3.13: Average autocorrelation function in three cell types.

cells (MSCs) that have been shown to be essential for HSCs maintenance and function [9, 101, 118]. Intravital microscopy (IVM) was used to record time-lapse images, and the interactions between cells of interest and nestin-GFP+ cells were able to be studied within the mice calvarium bone marrow.

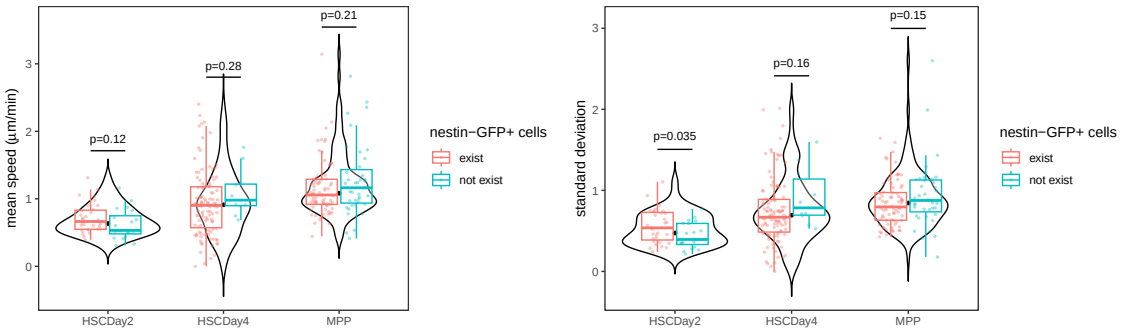
As explained in Section 3.4, after data cleaning, 213 cells had information about nestin-GFP+ cells. Notably, all HSC cells observed on the fourth day after transplantation were in nestin-GFP+ reporter mice. This suggested that the HSCDay4 cells lacking nestin-GFP+ information were located too far from nestin-GFP+ cells, resulting in their absence within the field of imaging view.

Figure 3.14 compared cell movement behaviours in the field of view with and without nestin-GFP+ cells. Permutation tests were conducted to statistically challenge the null hypothesis that the mean values of cell movement behaviours (mean speed, standard deviation, arrest coefficient and confinement index) are independent of whether the presence of nestin-GFP+ cells in the field of view or not. The two-tailed p-values for each test, estimated by the Monte-Carlo approximation with 250,000 test statistics, were reported in the figure, at the Bonferroni-corrected significance level $p < 0.05/3 = 0.0167$.

From permutation test results, there was no evidence to reject the null hypothesis that HSCDay2, MPP, and HSCDay4 have different mean speeds, standard deviations, and arrest coefficients, regardless of whether the cells were within a field of view containing nestin-GFP+ cells or not, as shown in Figures 3.14(a), 3.14(b), and 3.14(c). This showed that their movement speed has not changed due to the presence of nestin-GFP+ cells.

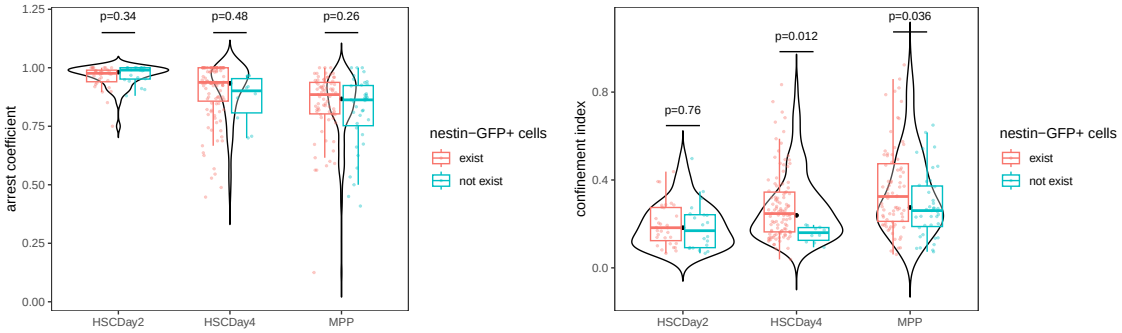
However, the confinement index for HSCDay4 displayed a statistically significant difference, with a p-value of 0.012 (< 0.0167). As mentioned before, all data at day 4 were collected from nestin-GFP+ reporter mice. This suggested that HSCs on the four-day post-transplantation exhibited significantly more nonlinear motion when they were far from nestin-GFP+ cells compared to when they were close to nestin-GFP+ cells. In contrast, day 2 cells (HSCDay2 and MPP cells) did not reject the null hypothesis, as shown in Figure 3.14(d). This showed that there was no difference in the linear movement of cells on the second day.

From this statistical analysis, it appeared that variables such as mean speed, standard deviation of speed, arrest coefficient, and confinement index were not affected by the presence of nestin-GFP+ cells for cells two days after transplantation. However, as time passed, cells close to nestin-GFP+ cells and those far away from them exhibited different



(a): Mean speed

(b): Standard deviation



(c): Arrest coefficient

(d): Confinement index

Figure 3.14: Comparison of cell movement behaviours in the field of view within and without nestin-GFP+ cells.

The red boxplots and points represent cells in environments with nestin-GFP+ cells, and the blue ones are the cells in the field of view without nestin-GFP+ cells. The black violin plots show the distribution of each variable in different cell types, and the black points in the middle of the violin plots are the median values.

movement behaviours four days after transplantation. This difference may be attributed to cell differentiation. The following analysis explores whether the differentiation has already been evident in the data from the second day.

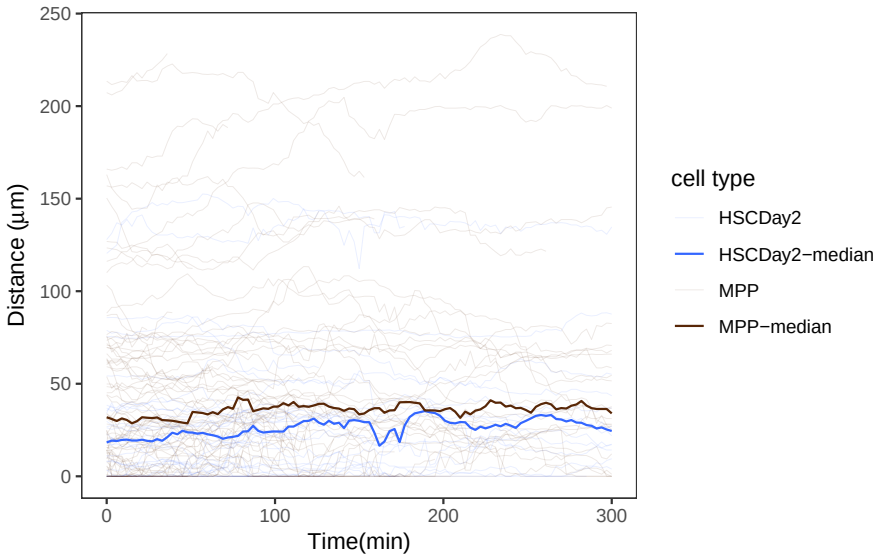


Figure 3.15: The distance from cells after 2 days transplantation to nestin-GFP+ cells.

Each light-blue line represents the distance of individual HSCDay2 cells to nestin-GFP+ cells over time, while the light-brown lines correspond to MPP. The blue line indicates the median distance for all HSCDay2 cells to nestin-GFP+ cells, and the brown line represents the median distance for all MPP.

As discussed in Section 3.4 about spatio-temporal data cleaning, information about nestin-GFP+ cells was mainly contained in two variables *Ncontact* and *nestin*. Twenty HSCDay4 cells contained ambiguous data records (e.g., "> 140" and "> 30"). In Figure 3.15, there was a representation of every track of HSC and MPP at day 2, with the median of all tracks for each population in a bold line. Over time, MPPs were generally located farther from nestin-GFP+ cells compared to HSCs.

To address this ambiguous information, a novel statistical variable was created the nestin arrest coefficient (NAC). Similar to the arrest coefficient, this new variable described whether the cell was moving closer to or farther away from nestin-GFP+ cells. The median distance from HSCDay2 to the nearest nestin-GFP+ cells was 18 μm at the initial time, while for MPP cells, it was 32 μm. The average initial distance to nestin-GFP+ cells

across both cell types was 25 μm. Furthermore, the NAC was defined as the percentage of time that the nearest distance between the nestin-GFP+ cell to the cell was less than 25 μm.

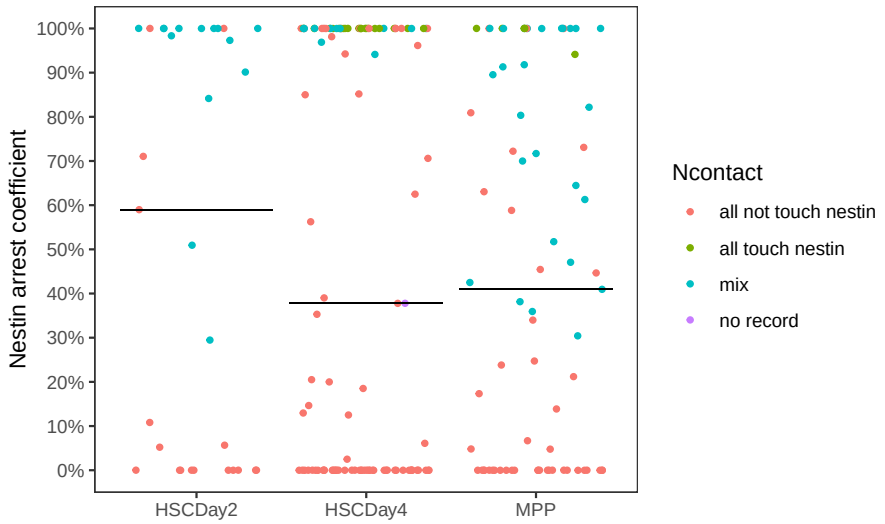


Figure 3.16: Nestin arrest coefficient.

The point plot shows the nestin arrest coefficient for different cell types, with colour indicating '*Ncontact*' status. The black line shows the median nestin arrest coefficient of each cell type.

Figure 3.16 showed the NAC for different cell types and included the variable *Ncontact* which showed whether a cell contacted the nestin-GFP+ cells. The black line in the figure represented the median value of the NAC for each cell type. The median of HSCDay2 was much higher than HSCDay4 and MPP, which revealed that most HSCDay2 have a preference for being located near nestin-GFP+ cells. The green dots in the figure represented cells that were always directly touching nestin-GFP+ cells, while blue dots indicated cells that were sometimes in contact. Both groups showed higher NAC values compared to cells that were never in direct contact with nestin-GFP+ cells (red).

A permutation test was applied to challenge the null hypothesis that the mean value of the nestin arrest coefficient is independent of cell types, and Bonferroni correction was also employed to get the Bonferroni-corrected p-value threshold of $p < 0.05/3 = 0.0167$. The two-tailed p-value was estimated from a Monte Carlo approximation with 250,000 statistics. The p-value between Day2 cells (HSCDay2 and MPP) was 0.46, between HSCs

on different days was 0.75, and between HSCDay4 and MPP was 0.7, providing no evidence to reject the null hypothesis. Unlike other variables, NAC did not differ significantly between cell types.

However, from Figure 3.16, MPPs had more cells with NAC between 0 and 1 compared to HSC cells. About 40% of HSC cells were closer to nestin-GFP+ cells and had higher NAC, and about 40% of HSCs were farther away from nestin-GFP+ cells and had NAC close to 0. Compared with HSCs, the proportion of MPPs located near nestin-GFP+ cells was smaller (about 20%).

Mesenchymal stromal cells (MSCs) secrete many relevant factors impacting the behaviour of haematopoietic cells, including HSCs [31]. This also raised the question of whether the varying distances of HSCs from nestin-GFP+ cells two days after transplantation were related to their movement behaviours.

Figure 3.17 compared each cell's mean speed when located within 25 μm of nestin-GFP+ cells versus when located farther than 25 μm. The x-axis represented the mean speed of cells when within 25 μm of nestin-GFP+ cells. For cells with NAC of 1, their mean speed when located beyond 25 μm was set as 0. Conversely, when NAC was 0, it indicated that the cell was always farther than 25 μm from the nestin-GFP+ cells, and their average speed within 25 μm was defined as 0. For cells (empty circle) with an NAC between 0 and 1, the x-axis showed the average speed within 25 μm, while the y-axis showed the average speed outside 25 μm.

The figure showed that blue points, representing HSCDay2, located in the lower-left corner indicated that HSCDay2 cells had a lower mean speed than HSCDay4 and MPP both close to the nestin-GFP+ cells or far away from them. For cells (empty circles in the figure) with an NAC greater than 0 but less than 1, the cell was sometimes close to nestin-GFP+ and sometimes far away. The mean speed of all empty circles was positively correlated when close or away from the nestin-GFP+ cells, with a Pearson correlation coefficient of 0.44. It suggested no evidence that proximity to nestin-GFP+ cells affected the speed of cells.

The foregoing statistical exploratory data analysis revealed that four days after transplantation, HSC cells exhibited significant differences in their linear movement depending on whether they were close to nestin-GFP+ cells or not. Although no statistically significant evidence in the movement behaviour of HSCs and MPPs was observed on the second day after transplantation, HSCs tended to be closer to nestin-GFP+ cells than MPPs.

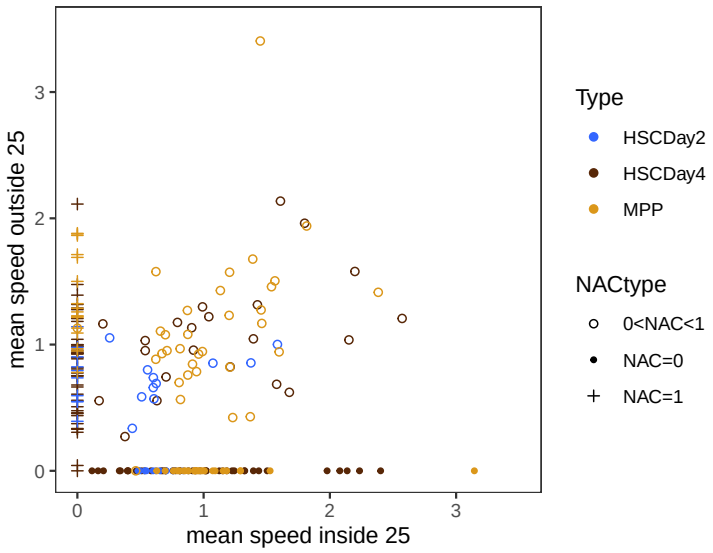


Figure 3.17: Each cell mean speed around nestin-GFP+ cells and far away the nestin-GFP+ cells.

The point plot shows the mean speed for each cell when they are close to the nestin-GFP+ cells or far away. The different colours report the cell types and the different shapes show the nestin arrest coefficient types. The x-axis indicates the cell mean speed when the cell is close to the nestin-GFP+ cells and the y-axis is the cell mean speed when the cell is far away to the nestin-GFP+ cells.

Additionally, HSCs appeared to be divided into two groups based on their NAC. Next, unsupervised clustering and biological marker analyses were used to investigate whether cell movement behaviour influences cell differentiation.

3.6 Clustering of cell tracks

From the previous spatio-temporal exploratory analysis of cell movements, there was no evidence to suggest that the mean speed, standard deviation, arrest coefficient, confinement index and nestin arrest coefficient differed between HSCDay2 and MPP cells. The uniform manifold approximation and projection for dimension reduction (UMAP) was used to cluster the cells on the second day after transplantation to capture cell similarities in spatio-temporal data [99].

Figure 3.18 presented the UMAP clustering result. It suggested that HSCDay2 could be divided into two subgroups, the top subgroup and the bottom subgroup. The points in Figure 3.19(a) provided the information for each variable, and the boxplots in Figure

3.19(b) showed the statistical summary of HSCDay2 movement behaviour in the top and bottom subgroups identified by UMAP.

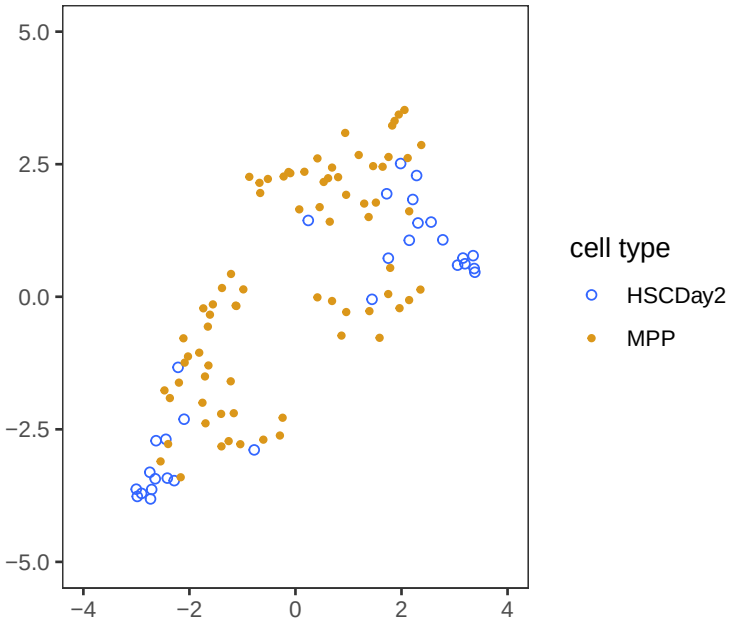
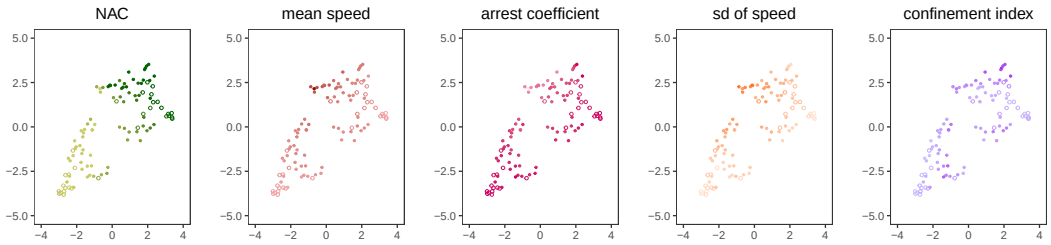


Figure 3.18: UMAP 2D visualisation.

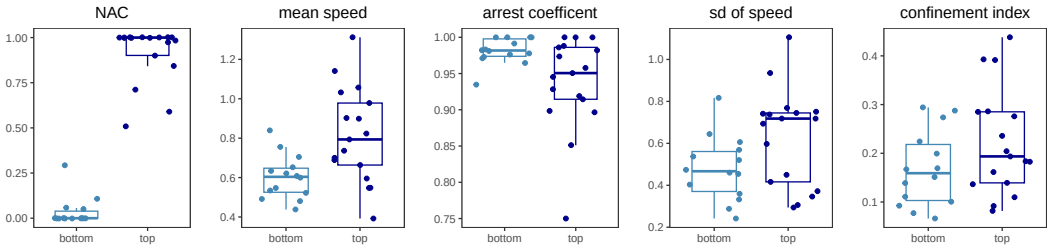
Two-dimensional projection visualisation show clustering results for day 2 data using UMAP. The yellow solid points show the MPP cells and each blue represents the HSCDay2 cell.

A permutation test was employed to statistically challenge differences in NAC between the bottom HSCDay2 subgroup and the top HSCDay2 subgroup. The null hypothesis is that the NAC is independent of whether the HSCDay2 cell belongs to the top or bottom UMAP subgroups. The two-tail p-valued was $p = 4 \times 10^{-6}$, which was estimated by Monte Carlo approximation with 250,000 test statistics, resulting in the null hypothesis being rejected. Although previous exploratory analysis did not show heterogeneity in NAC between different cell types, UMAP suggested the heterogeneity within the HSCDay2 population, which was strongly associated with whether cells maintained close proximity to nestin-GFP+ cells or not.

Permutation tests were also used to challenge the difference in mean speed, arrest coefficient, stand deviation of speed and confinement index between the bottom HSCDay2



(a): Cell movement behaviours visualisations.



(b): Subgroup visualisations.

Figure 3.19: Cell movement behaviours visualisations.

(a) The point plots indicate the movement behaviours for each cell in the UMAP result. The darker the point, the higher value of the variables for the cell. The empty dots represent the HSCDay2 and the solid dots are MPP. (b) The boxplots include statistical summaries for the bottom subgroup and top subgroup of HSCDay2 from UMAP clustering.

group and the top group. The two-tail p-values were 0.005, 0.013, 0.0046 and 0.1 for mean speed, arrest coefficient, stand deviation of speed and condiment index respectively, calculated by 250,000 Monte Carlo approximation. Combined with the visualisation, the subgroup of HSCDay2 cells located at the top right of the UMAP clustering had a significantly higher NAC and displayed more active, linear movement compared to the bottom subgroup, because the top group had a higher mean speed, a greater standard deviation of speed, a larger confinement index, and a lower arrest coefficient, indicating more dynamic behaviour. In contrast, the bottom group, located further from nestin-GFP+ cells, exhibited less active movement.

This clustering result suggested that two subgroups of HSCDay2 exhibited statistically different behaviours. One group actively approached the nestin-GFP+ cells, while the other did not. This information suggested that there was heterogeneity within this HSCDay2 population, with some cells might be in different stages of differentiation.

Following the report highlighting that the data suggested two groups, my collaborators at Imperial College London decided to reexamine a subset of the data. Once the imaging was completed, they sacrificed some of the mice, fixed the tissue, and stained for clusters of differentiation markers to determine the phenotypes of nine of the transplanted cells at the experiment's conclusion. The stained markers included a lineage cocktail for terminally differentiated cells, CD150, CD48, and CD41, which are well-established for identifying HSCs. This allowed them to confirm that, as indicated by the tracking data, the cell population exhibited phenotypic heterogeneity. Some cells matched the primitive HSCs phenotype (Lin-, CD150+, CD48-), while others expressed lineage markers, suggesting they had already begun differentiating.

Because it is not possible to determine differentiation status from IVM, this post-imaging staining process was crucial for validating the predictions made from the IVM data. Due to the complexity of this work, only nine cells were successfully analysed. Figure 3.20 showed the nine cell phenotype results placed, overlaid on the UMAP clustering analysis to confirm the relationship between cell behaviour and differentiation state. For the nine cells analysed, five cells still appeared as HSCs phenotype and represented by blue circles in Figure 3.20. The remaining four cells had differentiated into other cell types, and they were in yellow in the figure. At the upper right subgroup of the UMAP figure, four HSCs appeared, while cells in the lower right subgroup had differentiated into other cells. This means the UMAP clustering was largely correct, providing evidence in support of the inference.

The biological differentiation results combined with the UMAP clustering result suggest that cells remaining in the HSCs state tend to stay closer to nestin-GFP+ cells and exhibit more active behaviour. In contrast, as cells differentiate into other types, they might move away from nestin-GFP+ cells and show less active movement. These findings suggest that cells in distinct differentiation states can be classified based on their heterogeneity behaviours, highlighting the potential to use cell movement patterns as indicators of their developmental status.

Based on the analysis of stem cell motility on the second day, HSCs transplanted four days earlier were predicted using the UMAP. Figure 3.21 showed the results. It can be seen that the cells of HSCDay4 do not form obvious clusters. Some of them are in the upper right corner, overlapping with the HSCDay2 top subgroup, some are close to HSCDay2 in the lower left subgroup, and the rest align more closely with MPPs. This prediction result showed that the cell differentiation of HSCDay4 was diverse. Compared with the differentiation results on the second day, a larger proportion of HSCs transplanted four

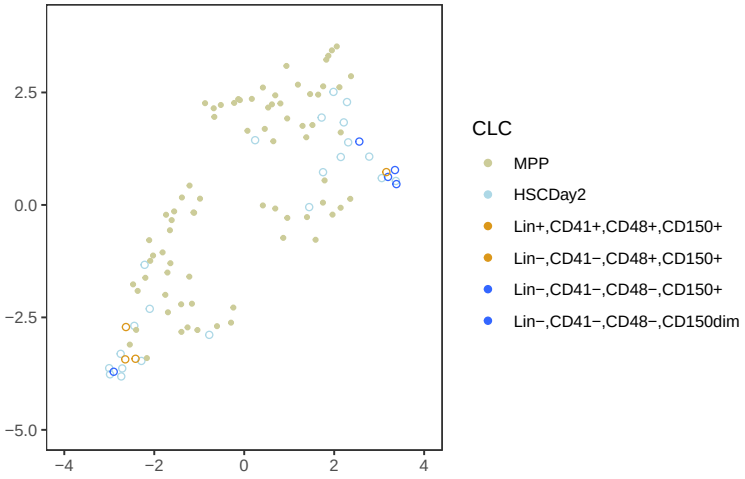


Figure 3.20: Biological markers result.

Two-dimensional projection visualisation displays the UMAP clustering results for Day 2 data. Nine immunostained cells are highlighted and distinguished by colour based on their cell phenotype.

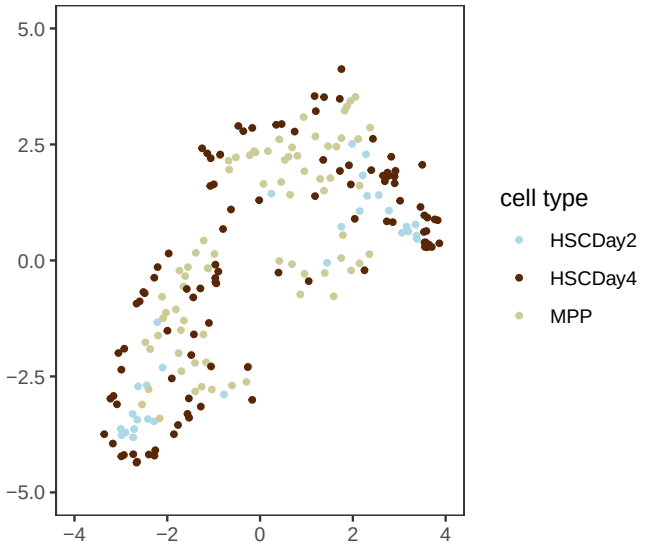


Figure 3.21: HSCDay4 prediction.

Two-dimensional projection visualisation displays UMAP prediction results for day 4 data. Light yellow points represent MPP, light yellow points represent HSCDay4, and predicted HSCDay4 are marked in brown.

days ago have differentiated into other cell types. Some of them are more similar to MPPs, and some cells may still be HSCs.

3.7 Discussion

In this chapter, extensive spatio-temporal data were collected from *in vivo* time-lapse microscopy by collaborators in Prof. Lo Celso's lab at Imperial College London. This dataset presented significant challenges due to the complexity of obtaining high-resolution images of live cells within the bone marrow of living mice. My main contribution was to design and implement a robust statistical pipeline for cleaning and analysing this complex dataset. It included format checks, completeness checks, identifying noise and data transformation to ensure high-quality reliable data. My other main contribution was to apply machine learning methods to explore spatio-temporal cell type distribution. This research characterises the unique movement and localisation behaviour of HSPC two days and four days after transplantation. These steps could be reproduced in other high-dimensional spatio-temporal datasets.

Previous studies have employed transcriptome sequencing and clustering methods to explore the differentiation process of HSPC in bone marrow, revealing that cells differentiate into various cell types and perform distinct functions at different time points following transplantation [38, 54, 72]. These studies however lack information on the localisations and behaviours of the cell examined. Furthermore, the bone marrow niche is critical for understanding the self-renewal and multipotency of HSCs [117], and the nestin-GFP+ cells in niche influenced HSCs behaviour by providing varying cytokine inputs necessary for HSCs maintenance, originating from distinct perivascular niches [117, 119]. While 3D *in vivo* imaging offers the opportunity to study the bone marrow environment, obtaining useful spatio-temporal data, especially over extended periods, has been complex and technically demanding.

Although the significance of the bone marrow niche has been recognised, no prior study has definitively demonstrated how nestin-GFP+ cells influence the precise spatial proximity of HSPC. This gap is due to the challenges in capturing *in vivo* time-lapse data and the complexity of analysing spatio-temporal data associated with cell dynamics. The statistical pipeline developed in this chapter was crucial in addressing these challenges, following the collection of the data by our collaborators.

The spatio-temporal data cleaning workflow was designed based on the complexity of high-dimensional data acquired by *in vivo* time-lapse microscopy. There were obstacles

that analysing the data about tracking HSPC in the dynamic bone marrow environment, including the presence of ambiguous records, data noise, and hard in accurately identifying cell positions relative to nestin-GFP+ cells. The spatio-temporal data cleaning workflow in this chapter involved not only standard preprocessing techniques, but also the development of new methods to handle and correct these ambiguities. For example, combined with visualisation, creating a new biological variable 'nestin arrest coefficient' accurately quantified the relationship between cell movement and proximity to nestin-GFP+ cells. This data cleaning workflow ensured that the dataset was robust and retained reliable biological meaning.

The spatio-temporal statistical analysis revealed significant differences in the movement behaviour of HSC four days after transplantation, depending on their proximity to nestin-GFP+ cells. Furthermore, the unsupervised learning method revealed the heterogeneity within the HSCs two days after transplantation, and discovered distinct subpopulations associated with nestin-GFP+ cells. The combination of immunophenotypic results with clustering results provided biological evidence to confirm that cellular behaviour was closely linked to differentiation status. It is a novel contribution in the cell type distribution analysis, demonstrating that cellular differentiation status can be inferred from movement patterns.

The spatio-temporal statistical analysis pipeline developed in this chapter has broad implications for haematopoietic stem progenitor cell dynamics. By combining spatio-temporal data cleaning with advanced statistical techniques, this research provides a framework for future studies to explore the dynamic behaviours of cells *in vivo*. The ability to link cell movement with differentiation states opens up new possibilities for understanding how the bone marrow environment influences haematopoietic stem cell fate decisions. Moving forward, the pipeline could be extended to other cell types and biological contexts, providing deeper insights into the spatio-temporal dynamics of cell behaviour *in vivo*.

4

Statistical analysis of temporal generational distribution in cells

4.1 Abstract

This chapter introduces a novel method called DivisionCounter, developed in collaboration with Dr. Leïla Perié's lab at Curie Institute, for analysing the long-term generational distribution of cells *in vivo*. A version of this work can be found online as a pre-print [65]. All biological experiments were designed and conducted by Dr. Perié lab, while my contribution was designing and implementing the statistical framework based on the data collected by our collaborators. This framework is well-suited for long-term cell data analysis.

Cell division drives multicellular growth and its dysregulation can lead to disease. While approximately 44 divisions are required to generate the 10^{13} cells in the human body, current methods are limited to counting 10 cell divisions *in vivo* across various mammalian cell types. DivisionCounter, the method introduced here, is based on a novel design. Mathematical results establish that if each cell in a population is equipped with a method to toss a random coin that, at each division, determines if offspring express a fluorescent protein, then knowing the likelihood of the change along with the proportion of label positive cells is sufficient to infer the average division count of the population. As described in this chapter, this approach was realised in practice by Dr Perié's Lab, through the construction of a genetic construct consisting of the code for a coin-toss mechanism and the expression of a fluorescent protein that is retrovirally loaded into cells, resulting in the DivisionCounter. It enables the estimation of cell division numbers *in vivo* over a large division range of up to 70 divisions using a simple fluorescence readout. Extensive experiments conducted in the Perié Lab and analysed with the statistical framework described in this chapter have demonstrated that the method can accurately measure

the average cell division in both *in vitro* and *in vivo*. Its use revealed that leukaemia tumour division rates are independent of organs' specific microenvironments and CAR-T cell treatment, providing an estimate of tumour death rate *in vivo*. The DivisionCounter method holds the unique potential to quantify the contributions of cell division, death, and migration to the growth of healthy and pathological mammalian tissues.

4.2 Introduction

Cell proliferation drives the growth of multicellular organisms during embryogenesis, tissue homeostasis, and regeneration after damage, and its dysregulation can lead to diseases such as cancer. It is well established that: i) cells are limited in the number of times they can divide, known as the Hayflick limit, of 40-60 divisions [63]; ii) cells accumulate genetic mutations during division, which can eventually lead to oncogenesis [59]; and iii) the memory of previous divisions regulate the decision of stem cells to differentiate [14]. In contrast, existing research still knows little about how many cell divisions are required to produce and maintain each organ in the body, or how cell divisions are regulated mechanically. This lack of knowledge about tissue development in healthy or pathological processes is largely due to a lack of high-resolution tools for counselling cell divisions *in vivo*.

Currently, available methods for quantifying the number of cell divisions in mammalian cells are limited in resolution and biological applications. One significant issue is that, while cell counting is often used as a proxy for cell division, it is unable to deconvolve cell death from cell division and so it measures the net cell population doubling rather than quantifying the number of cell divisions [34]. Live microscopy of cells can trace at most a few divisions in mammalian cells *in vivo* or tens of divisions *in vitro*, relying on direct observation of divisions [42]. Division diluting dyes, such as Carboxyfluorescein Diacetate Succinimidyl Ester (CFSE) and Cell Trace Violet (CTV), use the dilution of a dye by two at each division to infer division counts *in vitro* and *in vivo*, which has been shown in Figure 4.1 [94, 122]. However, these dyes are limited to counting 10 divisions at best due to the geometric dilution of the dye with division. Other fluorescence dilution–based methods have a resolution of 5 or 6 divisions [124]. Lengthier inference of division number has been realized by quantifying telomere lengths, but this method is restricted to a few specific cell types without telomerase activity. Therefore, there is a need for additional methods that can quantify more than 10 cell divisions in a broad range of mammalian cell types.

A novel methodology, the DivisionCounter, is presented in this chapter. The Division-

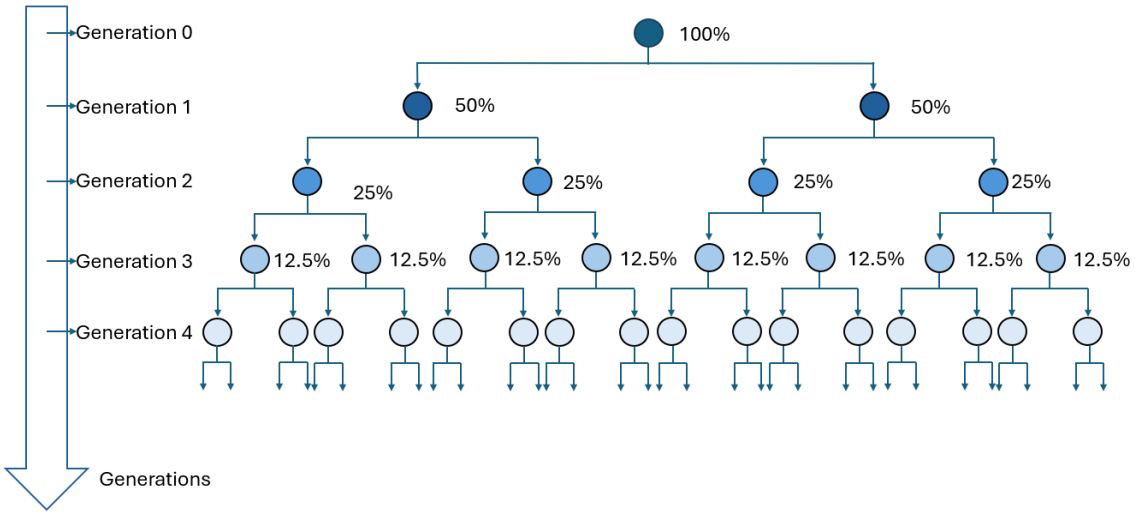


Figure 4.1: Cell tracker dye.

When cells divide, the fluorescence intensity is halved [95].

Counter can accurately estimate the number of divisions a population of cells has undergone up to more than 70, using a simple and low-cost fluorescence readout, that builds on previous qualitative recorders of cell proliferation. The DivisionCounter method consists of a lentiviral construct and a mathematical formula. When used together, they enable the estimation of the average division count of a cell population; that is, the mean of the division counts of each cell in the cell population of interest. The mathematical principles that ultimately led to the DivisionCounter were first published in 2016 [148]. It was mathematically established that the average division of a cell population could be inferred simply by tracking label performance in a cell population that had a small likelihood of label loss at each division event. The only conditions were that cells should all be initially labelled, lose their label expression at each cell division with a fixed probability, and never reacquire it. Implementing this approach in a biological system was achieved in 2022 [17]. The approach there, however, was qualitative rather than quantitative as the irreversible requirement of the genetic construct introduced significant experimental challenges. As a result, the mathematics of a system without that constraint were examined.

The dynamics of a more involved three-state system based around the inherent behaviour of the random per-division driver, microsatellites (MS), led to a new formula, which was called the DivisionCounter formula. In the first iteration of the mathematics, the loss is irreversible [43, 100, 148]. This system was experimentally implemented in an MS-driven system [17]. However, the results reported were qualitative. In the second iteration of the mathematics, which is developed in Section 4.2.2, the loss is no longer irreversible and

the math explicitly models MS mutation dynamics as they will drive the loss or gain of fluorescence. That result is the DivisionCounter formula used in this chapter.

4.2.1 Lentiviral construct

In search of a division-linked mutative element to build the DivisionCounter, MS was decided to use due to their low mutation rate that the likelihood per division was lower than 10^{-2} [81, 93]. MS mutations indeed generally occur by the addition or deletion of an entire nucleotide motif, which in this case is a single guanine. Therefore, by placing this MS between the DivisionCounter construct's start codon and the mScarlet gene, MS mutation would cause the addition or deletion of one nucleotide, therefore altering the DNA sequence, and more importantly shifting the transgene mRNA translation reading frame by +1 or -1 before the mScarlet gene thus causing improper mScarlet translation leading to mScarlet expression stoppage, which has been shown in Figure 4.2. Hence, the DivisionCounter construct label expression would either be on or off, however, the underlying mechanism of this label expression alteration comes from a three-state system, dictated by the biology of mRNA translation into protein depending on RNA sequences of nucleotides triplets, known as codons. While the original mathematical analysis relied on the irreversible label loss [148]. Here, a new three-state DivisionCounter formula was derived that takes into account the biological specificity of reversibility in this experimental system.

4.2.2 Derivation of the DivisionCounter formula

The first step in developing the DivisionCounter formula arises from the analysis of a model of a three-state system that tracks changes in mRNA translation reading frame states and its link to division-induced MS frameshift mutation. The three-state system was modelled as a Markov chain that captures the reading frame state of a given cell through divisions amongst the three possible mRNA reading frames, namely in the frame, +1 frameshift, and −1 frameshift. All cells will be assumed to start in the "in frame" state and, at each cell division, will have a cumulative frameshift mutation probability, ρ, to move to +1 reading frame after a guanine addition or to move to a −1 reading frame after a guanine deletion. If the cell is in state 1 and experiences a −1 mutation, it moves to state 3. Similarly, if a cell is in state 3 and experiences a +1 mutation, it moves to state 1. In this way, the state records the in- and out-of-frame state of the cell.

Let q be the probability of moving from the current state to the +1 reading frame state, and also to the −1 reading frame state, so that the probability of a mutation in either direction was $\rho = 2q$. This means a cell will have a probability to remain in its current

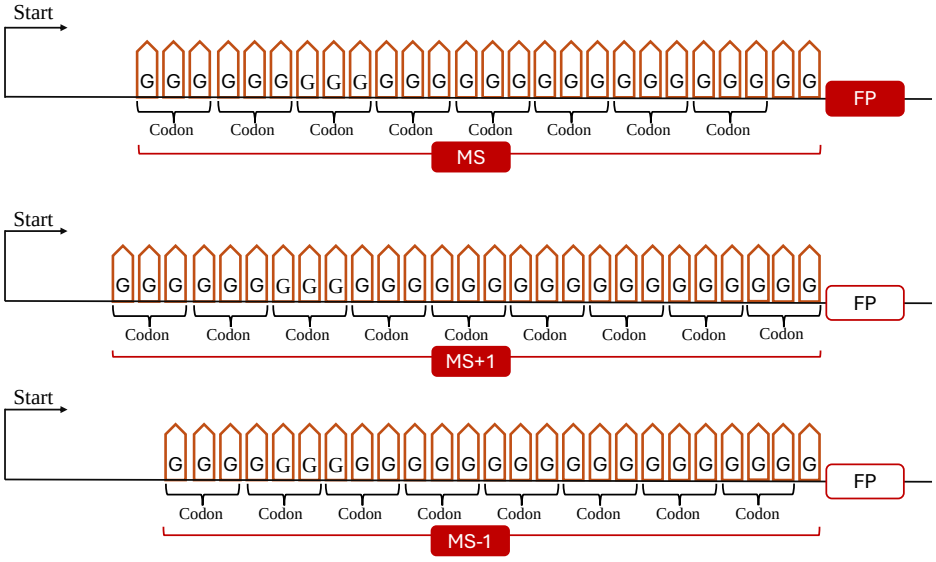


Figure 4.2: Schematic of the DivisionCounter construct.

The top panel shows the DivisionCounter construct, which contains two essential features a microsatellite (MS) and a fluorescence protein (FP) in-frame, in an open reading frame, all packaged in a lentivirus. After infection, the cells that are FP-positive are selected to start the DivisionCounter experiment. With each cell division, during DNA duplication, each cell has a low likelihood to undergo polymerase slippage on its MS sequence that will result in altering the MS length, pushing the FP out of the open reading frame, as illustrated in the middle and bottom panel. Reversibly, the MS+1 and MS-1 can subsequently come back to MS and the open reading frame during later cell divisions due to new slippage events.

state equal to one minus to probability to undergo either frameshift mutations, which is equal to $1 - 2q$. In mathematical terms, analysis of this system relies on the following Lemma.

Lemma 4.2.1. *:*
Let

$$\mathbb{P} = \begin{pmatrix} 1-2q & q & q \\ q & 1-2q & q \\ q & q & 1-2q \end{pmatrix}.$$

Then, for any $i \geq 1$*,*

$$\mathbb{P}^i = \begin{pmatrix} D_i & O_i & O_i \\ O_i & D_i & O_i \\ O_i & O_i & D_i \end{pmatrix}. \tag{4.1}$$

With $D_i = 1/3 + 2/3(1-3q)^i$ *and* $O_i = 1/3 - 1/3(1-3q)^i$*.*

Proof. Due to symmetries, $\mathbb{P}^i$ has the form in Equation 4.1 with $D_1 = 1-2q$ and $O_1 = q$. As $\mathbb{P}^i = \mathbb{P} \cdot \mathbb{P}^{i-1}$, $D_i = (1-2q)D_{n-1} + 2qO_{n-1}$ and $O_i = (1-q)O_{n-1} + qD_{n-1}$ were obtain. Subtracting the second equation from the first gives $D_i - O_i = (1-3q)(D_{i-1} - O_{i-1}) = (1-3q)^i$, for all $i \geq 1$. Hence, using one of the original inductive relations can be rewritten as $D_i = D_{i-1} - 2q(1-3q)^{i-1} = 1/3 + 2/3(1-3q)^i$ and the statement of the lemma follows. □

Corollary 4.2.1.1. *As can be seen from the expressions for* D_i *and* O_i*, the steady-state equilibrium is to have* $1/3$ *of cells be in the reading frame and* $2/3$ *of cells be out of the reading frame. This is referred to be the plateau.*

Now that a three-state reading frame system was modelled, it used to drive an average division estimator formula, $\widehat{G}(t)$, of a given cell population, $Z(t)$, at time t, based on its percentage label expression, meaning the percentage of cells in the 'in the reading frame' state, $Z_1(t)/Z(t)$, where $Z_1(t)$ is the number of cells in 'in the reading frame' state at time t. In order to do so, the cell population $Z(t)$ was defined as the sum of all cells in each reading frame state (1, 2, and 3) at time t, and was characterised nu a vector of proportions of cells with different initial states, $\overline{\pi} = (\pi_1, \pi_2, \pi_3)$, setting to $\pi_1 = 1$ and $\pi_2 = \pi_3 = 0$ as all cells will start in 'in the reading frame' state as label positive cells. This resulting formula is referred to be the DivisionCounter formula 4.2.

Theorem 4.2.2. *If $\pi_1 = 1$ and $q = \rho/2$, then almost surely*

$$\widehat{G}(t) = \frac{1}{\log(1-\frac{3\rho}{2})} \log\left(\frac{3\mathbb{E}(\frac{Z_1(t)}{Z(t)}) - 1}{2}\right) . \tag{4.2}$$

Proof. Let $Z(t) = Z_1(t) + Z_2(t) + Z_3(t)$ the total number of cells at time t. Then,

$$\mathbb{E}\left(\frac{Z_j(t)}{Z(t)}|Z(t)\right) = \mathbb{E}\left(\frac{1}{Z(t)}(\sum_{k=1}^{Z(t)} I_{k,t})\right) = \frac{1}{Z(t)}\sum_{k=1}^{Z(t)} \boldsymbol{p}_{k,t} = \frac{1}{Z(t)}\sum_{k=1}^{Z(t)} \left(\overline{\pi}\boldsymbol{p}^{G_k(t)}\right)_j ,$$

where $I_{k,t}$ is the indicator function that cell k is of type j and $\boldsymbol{p}_{k_t}$ is the probability that cell k is of type j. Then, taking expectations,

$$\begin{aligned}
\mathbb{E}\left(\frac{Z_j(t)}{Z(t)}\right) &= \mathbb{E}\left((\overline{\pi}\boldsymbol{p}^{G(t)})_j\right) \\
&= \mathbb{E}\left(\pi_j(\frac{1}{3} + \frac{2}{3}(1-3q)^{G(t)})\right) + (1-\pi_j)\left(\frac{1}{3} - \frac{1}{3}(1-3q)^{G(t)}\right) \\
&= \frac{1}{3} + (\pi_j - \frac{1}{3})\mathbb{E}\left((1-3q)^{G(t)}\right) ,
\end{aligned}$$

or

$$\mathbb{E}\left((1-3q)^{G(t)}\right) = \frac{3\mathbb{E}(\frac{Z_j(t)}{Z(t)}) - 1}{3\pi_j - 1} .$$

Hence, $\log \mathbb{E}(e^{sX}) = s\mathbb{E}(X) + o(s)$ for small s, one can use the estimator

$$\widehat{G}(t) = \frac{1}{\log(1-3q)} \log\left(\frac{3\mathbb{E}(\frac{Z_j(t)}{Z(t)}) - 1}{3\pi_j - 1}\right) ,$$

where for q is smaller than $1/3$, $\pi_1 = 1$ and $q = \rho/2$. □

Based on the DivisionCounter formula, the biological collaborators designed *in vitro* and *in vivo* experiments. The *in vitro* experiments were mainly used to confirm that DivisionCounter can count in a robust, quantitative manner and obtain the probability of mutation of specific cell types. The *in vivo* experiments were mainly divided into two parts. One part aims to measure the cancer cell division rate across different organs. The other part measures the *in vivo* tumour division rate during CAR-T cell treatment and gets the estimation of the tumour death rate. Both sets of experiments would necessitate careful statistical analysis to challenge scientifically meaningful hypotheses.

4.3 Methods

This section explains the dataset and the statistical methods used in this chapter. While some of these methods may be familiar to the reader, their inclusion ensures that the thesis is self-contained and comprehensive. All programming was performed using `R` (version `R` 4.3.2). The data analysed in this section consists of two parts, *in vitro* and *in vivo*, both of which are temporal datasets. The *in vitro* analysis begins with data cleaning, followed by the application of advanced statistical methods to estimate key parameters for DivisionCounter. The focus then shifts to the *in vivo* data, where robust statistical approaches assist collaborators in understanding cell behaviour from two perspectives: cell numbers and cell division counts.

4.3.1 Statistical inference of fluorescence loss probability

Dataset description

To validate the effectiveness of the DivisionCounter method, *in vitro* experiments were used to evaluate its performance in biological systems. The entire experiment was repeated several times independently, known as experimental replication, to assess variability across different biological samples. Technical replication was to check the consistency and accuracy of the measurements, and the same sample was measured multiple times using the same measurement method. The nonlinear mixed-effects model described the difference changes between each experimental measurement and the fixed effect of the experiment [87, 125].

In the original dataset collected by our collaborators from *in vitro* experiments, there are seven variables, and each variable has been explained in Table 4.1. After the format check and completeness check, all data with the percentage of mScarlet-positive less than 35% were excluded.

Variable	Detail	Description
experiment number	Integer	Biological experiment number
time point days	Integer	Data collection time
condition	Character	23G, 26G, Control
temperature	Numerical variable	33, 37
technical repeat	Integer	1,2,3
division count	Numerical variable	cell population doublings
percentage mScarlet pos	Numerical variable	%mScarlet cells

Table 4.1: Variables in the *in vitro* dataset to estimate the probability of mutation.

Statistical framework

Let $m_{t_j^k,i}$ represents the percentage of mScarlet-positive cells at jth time point, kth technical repeat, and ith experiment. Let $D_{t_j^k,i}$ is the cell population doublings at jth time point, kth technical repeat and ith experiment, where $1 \leq j \leq J$, $1 \leq k \leq K$, $1 \leq i \leq I$, and $j, k, i \in \mathbb{N}^+$ and $\mathbb{N}^+$ is the set of positive integer.

The mean percentage of mScarlet-positive cells for time j at experiment i is

$$\overline{m}_{t_j,i} = \frac{m_{t_j^1,i} + m_{t_j^2,i} + \cdots + m_{t_j^K,i}}{K} ,$$

and the mean cell population doubling for jth time and ith experiment is

$$\overline{D}_{t_j,i} = \frac{D_{t_j^1,i} + D_{t_j^2,i} + \cdots + D_{t_j^K,i}}{K}.$$

Based on Formula 4.2, a nonlinear mixed effects model was used to estimate the probability of mutation:

$$\overline{D}_{t_j,i} = \widehat{g}(\widehat{\rho}_i, \overline{m}_{t_j,i}) + \epsilon_{t_j,i} , \tag{4.3}$$

where $\overline{D}_{t_j,i}$ was the response vector, $\widehat{g}$ was a nonlinear function with parameter $\widehat{\rho}_i$ the predictor vector $\overline{m}_{t_j,i}$, and $1 \leq t_j \leq t_J$ and $1 \leq i \leq i$. The $\epsilon_{t_j,i}$ was the normal distribution error term [87, 116].

In detail, $\widehat{\rho}_i$ could be written as

$$\widehat{\rho}_i = A_i\beta^f + B_i\beta_i^r,\ \beta_i^r \sim N(0, \sigma^2 D),$$

where β^f was a vector that included the fixed effects parameters with length $\left|\beta^f\right|$, and β_i^r represented the $|\beta_i^r|$-dimensional random effects vector which only related with i (the ith experiment). The matrix A_i was for fixed effects and B_i was for random effects.

The following formula is the marginal density of the $\overline{D}$,

$$p(\overline{D}) = \int p(\overline{D}|\beta^r)p(\beta^r)d\beta^r.$$

Then, the nonlinear mixed-effects model parameters were estimated by the maximum likelihood with a penalized nonlinear least squares step and a linear mixed effect step for high speed and sensitivity [136]. It was implemented by `R` package `'nlme'` (version

3.1-163) to get the best estimate of the probability of the losing mScarlet-expression ρ [115].

The bootstrap method was used to evaluate 95% confidence intervals (CI). In detail, a bootstrap sample was created by sampling with replacement from the experiment number and technical repeat of the dataset and an estimate was obtained from the bootstrap sample via the nonlinear mixed effects model. This process was repeated 100,000 times to estimate the ordered bootstrap samples $\widetilde{\rho}$. The $\widetilde{\rho}_{(2.5\%)}$ and $\widetilde{\rho}_{(97.5\%)}$ represented the 2.5th and 97.5th percentiles of bootstrap distribution. Hence, the basic upper and lower bootstrap confidence intervals were determined as $[\widetilde{\rho}_{(2.5\%)}, \widetilde{\rho}_{(97.5\%)}]$ [143].

In order to evaluate the probability of losing mScarlet-expression cultured at 37°C and 33°C, the Pearson correlation coefficient was used [128]. Using Formula 4.2 and the best estimate ρ_1 from 37°C, the average division count were predicted in set $G_2 = \left\{\widehat{g}(\rho_1, m_{t_j^k,i})|1 \leq j \leq J, 1 \leq k \leq K, 1 \leq i \leq I, j, k, i \in \mathbb{N}^+\right\}$ at 33°C. The set D_2 includes all cell population doublings information at 33°C from the biological experiment. The Pearson correlation coefficient was calculated by $\mathrm{cov}(G_2, D_2)/(\sigma(G_2)\sigma(D_2))$, where 'cov' was the covariance, $\sigma(G_2)$ was the standard deviation of G_2, and $\sigma(D_2)$ was the standard deviation of D_2.

The range of the Pearson correlation coefficient is $[-1, 1]$ [128]. The closer it is to 1, the more correlated between the predicted division count G_2 and the cell population doublings D_2. It also means the probability of losing mScarlet-expression cultured at 37°C is closer to the probability of losing mScarlet-expression cultured at 33°C. It was performed by '`cor`' function with method '`pearson`' in `R`.

An ordinary least squares linear regression was used to get the division rate of the cell per day, because it offers an approach for modelling the linear relationship between the cell average division counts and time. The model is

$$G_{j,k,i} = G_0 + g \times t_{j,k,i} + \epsilon_{j,k,i} \ , \tag{4.4}$$

where the response variable $G_{j,k,i}$ represented the predicted average division count for the ith experiment, at the jth time point and the kth technical repeat. The intercept G_0, which corresponded to the average division count at day 0, was fixed at 0. The model coefficient was g, and the explanatory variable $t_{j,k,i}$ represented the time point for the ith experiment, jth time point, and kth technical repeat. The $\epsilon_{j,k,i}$ was the error term.

The residual sum of squares (RSS) was defined as

$$\mathrm{RSS}(g) = \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{i=1}^{I}(G_{j,k,i} - \widetilde{g}t_{j,k,i})^2 \ ,$$

such that the approximate minimum was found

$$\widetilde{g} \in \arg\min_{g} \mathrm{RSS}(g) \ ,$$

where $\widetilde{g}$ is the best estimate of division rate that minimizes $\mathrm{RSS}(g)$. The 95% confidence intervals were calculated by $[\widetilde{g}-\theta(g), \widetilde{g}+\theta(g)]$, where $\theta(g)$ was the t critical value multiply the standard deviation of division rate in the model. The t critical value was the cutoff point on the t distribution with 95% confidence level and degrees of freedom from the input dataset. This model and confidence intervals were implemented in `R` by '`lm`' function.

Using CTV, the median fluorescence intensity decreases as cell division increases. The set G_c included the average division count predicted with the best estimate of the probability of losing mScarlet-expression and $|G_c|$ was the number of elements in the G_c. The set F_c contained all median fluorescence intensity collected by CTV measurements, where $|G_c| = |F_c|$.

The rank of G_c and F_c were $R(G_c)$ and $R(F_c)$, respectively. The Spearman correlation coefficient was calculated using $\mathrm{cov}(R(G_c), R(F_c))/\left(\sigma(R(G_c))\sigma(R(F_c))\right)$, where 'cov' was the covariance of the rank data, $\sigma(R(G_c))$ was the standard deviation of $R(G_c)$, and $\sigma(R(F_c))$ was the standard deviation of $R(F_c)$ [151]. This calculation was performed using the '`cor`' function with method '`spearman`' in `R`.

4.3.2 Growth rate inference from cell numbers

Data description

In the original dataset collected by our collaborators from *in vivo* experiments, there are six variables, and each variable has been explained in Table 4.2. After the format check and completeness check, all data with cell numbers less than 455 were excluded. The reason for 455 is explained in Section 4.4.2.

Statistical framework

The set $V = \left\{(\varrho_{t_j^k,N,o_l}, \varrho_{t_j^k,G,o_l})|j = 1,2,\cdots,J, k = 1,2,\cdots,K, l = 1,2,\cdots,L\right\}$ includes cell counts collected by flow cytometry with counting beads for lth organ, where $\varrho_{t_j^k,N,o_l}$

Variable	Detail	Description
Mouse ID	Integer	Identifying unique mice
Day	Integer	Data collection time
Condition	Character	26G
Organ	Character	Bone Marrow, Liver, Lungs, Blood, Spleen
Panel used	Character	Biological information
Nalm6 Count	Numerical variable	Cell numbers

Table 4.2: Variables in the *in vivo* dataset for cell numbers.

showed the kth Nalm6 cell counts collected at time j in the 26N microsatellite in lth organ, and $\varrho_{t_j^k,G,o_l}$ represents the kth cell counts at time j ithe 26G microsatellite in lth organ.

In order to challenge the null hypothesis that the mean cell counts were independent of whether the mice in the 26N microsatellite or 26G microsatellite on the same day in the same organ, a permutation test was used [83].

The set $A_v = \left(\varrho_{t_j^1,N,o_l}, \varrho_{t_j^2,N,o_l}, \cdots, \varrho_{t_j^K,N,o_l}, \varrho_{t_j^1,G,o_l}, \varrho_{t_j^2,G,o_l}, \cdots, \varrho_{t_j^K,G,o_l}\right)$ was an ordered set, and it had $2 \times K$ elelemts. The first K elements included the 26N microsatellite cell counts at day j for organ l, and the last K cell counts collected in 26G microsatellite mice at day j for organ l. The real-valued statistic used under this null hypothesis is

$$M(A_v) = \frac{\sum_{k=1}^{K} \varrho_{t_j^1,N,o_l}}{K} - \frac{\sum_{k=1}^{K} \varrho_{t_j^1,G,o_l}}{K} .$$

A permutation π is created, and it reassigns labels to individual vectors. A recorded observation set is obtained $A_{\pi(v)} = \Big(\varrho_{\pi(t_j^1,N,o_l)}, \varrho_{\pi(t_j^2,N,o_l)}, \cdots, \varrho_{\pi(t_j^K,N,o_l)}, \varrho_{\pi(t_j^1,G,o_l)}, \varrho_{\pi(t_j^2,G,o_l)}, \cdots, \varrho_{\pi(t_j^K,G,o_l)}\Big)$. A collection of permutations Q can be characterised so that reordering $\left\{A_{\pi(v)}\right\}_{\pi \in Q}$ are equally likely under the null hypothesis. The empirical distribution of $M\left(A_{\pi(v)}\right)_{\pi \in Q}$ is used to compare with $M\left(A_v\right)$. Therefore, the Monte Carlo approximation with $|Q|$ test statistics was applied to estimate the left-tailed p-value and right-tailed p-value. Bonferroni corrections were performed. In detail, the left-tailed p-value is

$$p_l = \frac{1 + \sum_{j=1}^{|Q|} I_{M(A_v) \geq M(A_{\pi(v)_j})}}{|Q|},$$

and right-tailed p-value

$$p_l = \frac{1 + \sum_{j=1}^{|Q|} I_{M(A_v) \leq M(A_{\pi(v)_j})}}{|Q|},$$

where I is the indicator function.

Permutation tests and Monte Carlo approximation were also used to statistically challenge differences in luciferase measurements, which collected cancer cell metastasis data, between 26N microsatellite and 26G microsatellite on the same day. The null hypothesis was the mean of the luciferase measurements was independent of whether in the 26N microsatellite or in the 26G microsatellite on the same day. Bonferroni correction was performed.

To challenge the null hypothesis that the mean cell counts were independent of whether in different organs on the same day, permutation tests were used. The real statistic for the l_1th organ and the l_2th organ was

$$M(A_o) = \frac{\sum_{k=1}^{K} \varrho_{t_j^k, o_{l_1}}}{K} - \frac{\sum_{k=1}^{K} \varrho_{t_j^k, o_{l_2}}}{K} .$$

Then, Monte Carlo approximation was used to estimate the left-tailed and right-tailed p-values, and Bonferroni correction was used.

Assuming Nalm6 cell counts were no death *in vivo*, an exponential model was employed to estimate the growth rate of Nalm6 cell count

$$\varrho_{t_j^k} = \varrho_{t_0} \times 2^{\varpi \times t_j^k} + \epsilon_{t_j^k} ,$$

where $\varrho_{t_j^k}$ was the response variable, ϱ_{t_0} was the parameter, which was the initial value of the cell counts on day 0, ϖ was the growth rate parameter, t_j^k was the predictor variable, and $\epsilon_{t_j^k}$ was the error term for the kth observation at day j, $1 \leq j \leq J$, $1 \leq k \leq K$.

To achieve parameters ϱ_{t_0} and ϖ, minimizing the residual sums of squares (RSS) with Leveberg-Marquardt algorithm was performed due to the robustness [106]. It was implemented in `R` package '`minpack.lm`' (version 1.2-3) with function '`nlsLM`'. The RSS was defined as

$$\mathrm{RSS}(\varrho_{t_0}, \varpi) = \sum_{j=1}^{J} \sum_{k=1}^{K} \left(\varrho_{t_j^k} - \varrho_{t_0} \times 2^{\varpi \times t_j^k} \right) ,$$

such that the approximate minimum was found

$$(\varrho_{t_0}^*, \varpi^*) \in \arg\min_{\varrho_{t_0}, \varpi} \mathrm{RSS}(\varrho_{t_0}, \varpi) \ .$$

The 95% confidence interval (CI) was obtained from the bootstrap method. A bootstrap sample was designed by sampling with replacement from the time of the dataset. The growth rate was estimated by the exponential model from the bootstrap method. This action was iterated 100,000 times to estimate the ordered bootstrap samples $\widetilde{\varpi}$. The $\widetilde{\varpi}_{(2.5\%)}$ and $\widetilde{\varpi}_{(97.5\%)}$ represented the 2.5th and 97.5th percentiles of bootstrap distribution. Hence, the basic upper and lower bootstrap confidence intervals were determined as $[\widetilde{\varpi}_{(2.5\%)}, \widetilde{\varpi}_{(97.5\%)}]$ [143].

The cell counts on the day J were predicted by the first $J-1$ days cell counts employing the best-fit estimate ϖ^* of the exponential model. The permutation test was used to challenge the actual cell counts on day J and the predicted cell counts on day J. The null hypothesis for the permutation test was the mean number of cell counts at day J was independent of whether the cell counts were predicted or not.

The set $V_{J-1} = \left\{ \varrho_{t_j^k} | j = 1, 2, \cdots, J-1, k = 1, 2, \cdots, K \right\}$ included cell counts collected at the first $J-1$ days. The set $V_J^* = \left\{ \varrho_{t_j^k} \times 2^{\varpi^* \times (t_J - t_j)} | j = 1, 2, \cdots, J-1, k = 1, 2, \cdots, K \right\}$ was the predicted cell counts on day J, and there were $|V_J^*|$ elements in the set. The set $V_J = \left\{ \varrho_{t_J^k} | k = 1, 2, \cdots, K \right\}$ was the cell count collected by flow cytometry with counting beads and the set V_J includes $|V_J|$ elements. The real-valued statistic used under this null hypothesis is

$$M(A_J) = \frac{\sum_{k=1}^{K} \varrho_{t_J^k}}{|V_J|} - \frac{\sum_{k=1}^{K} \sum_{j=1}^{J-1} \varrho_{t_j^k} \times 2^{\varpi^* \times (t_J - t_j)}}{\left|V_J^*\right|} \ ,$$

A collection of permutations was created, and it reassigned labels to individual vectors. The empirical distribution was used to compare with $M(A_J)$. Therefore, the Monte Carlo approximation with test statistics was applied to estimate the left-tailed p-value and right-tailed p-value.

4.3.3 Division rate inference from DivisionCounter division counts

Data description

In the original dataset collected by our collaborators from *in vivo* experiments, there are seven variables, and each variable has been explained in Table 4.3. After the format check

and completeness check, all data with cell numbers less than 455 were excluded. In this dataset, a few day 0 *in vitro* data were added to get an estimate of the initial average division count.

Variable	Detail	Description
Mouse ID	Integer	Identifying unique mice
Day	Integer	Data collection time
Condition	Character	26G
Organ	Character	*In vitro*, Bone Marrow, Liver, Lungs, Blood, Spleen
Panel used	Character	Biological information
Nalm6 Count	Numerical variable	Cell number
%mSca	Numerical variable	%mScarlet cells

Table 4.3: Variables in the *in vivo* dataset for division counts.

Statistical framework

Using the probability of losing the mScarlet-expression *in vitro* as the parameter ρ_G to fit the Formula 4.3, average division counts *in vivo* for organ l were predicted by the percentage of mScarlet-positive cells with Formula 4.2 in the set $\widehat{G}_G = \left\{\widehat{G}(t_j^k, o_l) | 1 \leq j \leq J, 1 \leq k \leq K, 1 \leq l \leq L, j, k, l \in \mathbb{N}^+\right\}$.

A simple linear regression was employed like Formula 4.4. The output variable was the average division counts obtained with DivisionCounter the set $\widehat{G}_G$ and the input variable was the time (days post-injection). A bootstrap sample with 100,000 datasets was created to get 95% confidence interval by sampling with experiment number and technical repeat from the *in vitro* dataset, and time from the *in vivo* dataset.

Permutation tests and Bonferroni correction were used to statistically assess the average division counts obtained with DivisionCounter between different organs on the same day. The null hypothesis was that at day j the mean of the average division count obtained with DivisionCounter was independent of whether the data was collected from organ l_1 or organ l_2. The real statistic was

$$M(A_G) = \frac{\sum_{k=1}^{K} \widehat{G}(t_j^k, o_{l_1})}{K} - \frac{\sum_{k=1}^{K} \widehat{G}(t_j^k, o_{l_2})}{K} .$$

The left-tailed p-values and right-tailed p-values were calculated by Monte Carlo approximation with 500,000 sampling.

4.4 Results

4.4.1 Measurement and inference of cell divisions in vitro

To assess the performance of the DivisionCounter method in a biological system, our collaborators in Dr Perié's lab at Curie Institute infected HEK293T cells with either a lentivirus containing a 26-guanine (26G) microsatellite followed by the fluorescent protein (FP) mScarlet (herein, **26G-DivisionCounter**) or with a control lentivirus in which the microsatellite was replaced by a linker sequence of 26 nucleotides (**26N-construct**). Using a low multiplicity of infection, one copy was integrated on average of the DivisionCounter construct per cell, resulting in approximately 10% of the cells being infected. A clear expression of mScarlet was obtained for both conditions. After 3 to 7 days of culture, mScarlet-positive cells were sorted and grown in culture for 46 days, which has been shown in Figure 4.3. At each splitting, cells were counted, and a fraction of the cells was retained for a DivisionCounter readout, by recording the percentage of mScarlet-positive cells by flow cytometry. Each readout represents an independent acquisition from three independent experiments, each done in triplicates called 'technical repeat'. As a comparison control, cells were counted to compute cell population doubling, which is equivalent to the average division count in this culture condition for which there is no, or only minimal, cell death.

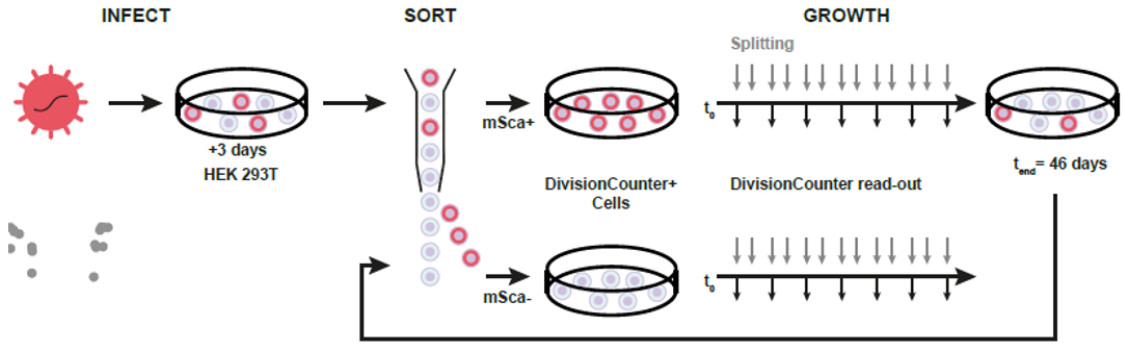


Figure 4.3: HEK293T cells *in vitro* experiment.

HEK293T cells were infected with the 26G- or 26N-DivisionCounter lentivirus at a low multiplicity of infection. After 3 days, mScarlet+ cells were sorted and cultured for 46 days. Once or twice a week when cells were split, cells were counted and analysed for their mScarlet expression by flow cytometry as a readout of the DivisionCounter method. Later, cells that became mScarlet-negative were also sorted, cultured and analysed for their mScarlet expression over time.

As predicted by the DivisionCounter method, the percentage of mScarlet-positive cells

infected with the **26G-DivisionCounter** decreased as the population proliferated, in contrast to cells with the **26N-construct** in Figure 4.4. The changes in the percentage of mScarlet-positive cells followed the predicted pattern, decreasing with the average cell population doubling and reaching a plateau at 35%, corresponding to the three-state equilibrium of the DivisionCounter method. This proved that the DivisionCounter formula was robust and reproducible across independent experiments.

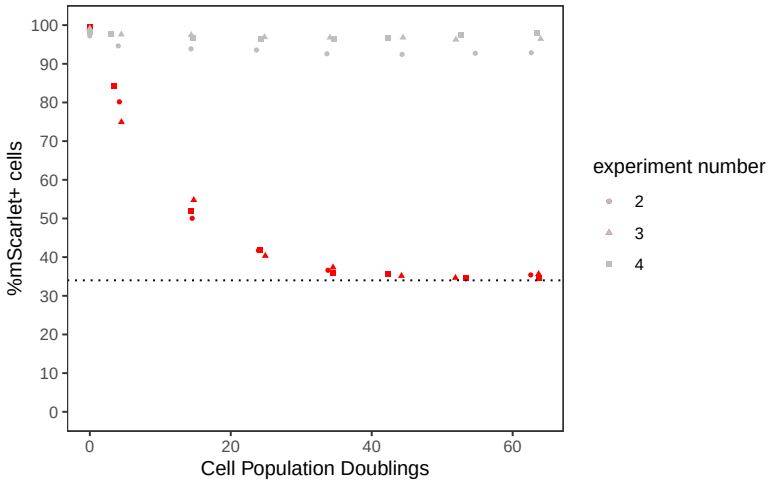


Figure 4.4: Three experiments data for 26G- and 26N-DivisionCounter.

Percentage of mScarlet+ cells over the average division counts measured using cell counts (population doublings) for **26G-DivisionCounter** (red) and **26N-DivisionCounter** (grey). Each point represents an independent acquisition, from three experiments done in triplicates, and the different shapes of the points show the different independent experiments. The black line is the plateau at 35%.

After demonstrating that this behaviour was robust and reproducible, we created and implemented a statistical framework to statistically explain the *in vitro* experiment performance. A nonlinear mixed-effects model was used to obtain ρ, the probability of loss of mScarlet-expression per division. All data showing that the percentage of mScarlet-positive cells was less than 35% were excluded before processing. The experiment number was treated as a random effect and the probability of the losing mScarlet-expression was a fixed effect. The bootstrap method was used to evaluate 95% confidence intervals (CI). For **26G-DivisionCounter**, the probability of loss of mScarlet-expression per division was $\rho = 0.047$ and CI $\in [0.044, 0.049]$, which has been shown in Figure 4.5. After obtaining the best estimate of ρ and its CI, substitute them into the DivsionCounter formula 4.2, and let $\mathbb{E}\left(Z_1(t)/Z(t)\right)$ be 35%. Hence, the maximum number of generations tracked

by DivisionCounter formula was obtained. The 26G-DivisionCounter counted up to 56 divisions, with a CI of 54 to 58 divisions. The detail of model fitting is seen in Section 4.3.1.

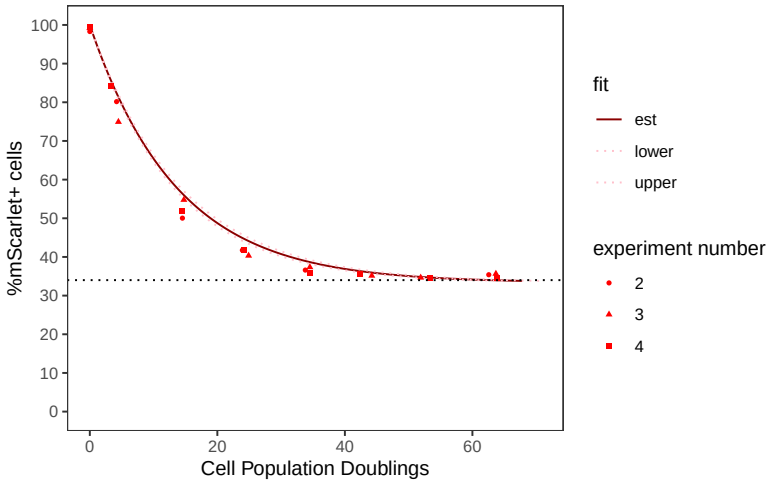


Figure 4.5: The probability of mutation for 26G-DivisionCounter.

Percentage of mScarlet-positive cells over the average division counts measured using cell counts (population doublings) for the **26G-DivisionCounter** (red). Each point represents an independent acquisition, from three experiments done in triplicates. The probability of losing mScarlet-expression ρ, was fitted by a nonlinear mixed-effects model. The dark red full line shows the best fit, and the light red dot lines show the lower and upper basic bootstrap bound of 95% CI, respectively. The black line is the plateau at 35%.

Then the robustness, versatility and general applicability of the DivisionCounter were explored by testing other microsatellite lengths and different cell lines. First, cells with a DivisionCounter construct bearing a 23G microsatellite rather than a 26G microsatellite were infected. Conducting the same experiment as before with infected HEK293T cells, all major behaviours of the DivisionCounter method were recapitulated, such as the percentage of mScarlet-positive cells decreased with the average cell population doublings and reached a plateau at 35% in Figure 4.6, validating the robustness of the DivisionCounter method. As the mutation rate of the guanine repeat–microsatellite increases with the number of guanine repeats, the 23G microsatellite was expected to mutate at a lower rate than the 26G microsatellite, thereby increasing the maximum division number that can be counted [77]. Figure 4.6 showed fitting of the probability of loss of mScarlet-expression for **23G-DivisionCounter**, the best estimate is 0.036 and the CI $\in [0.034, 0.038]$. The figure also shows more divisions were expected with the **23G-DivisionCounter** than

with the **26G-DivisionCounter**, reaching a maximum of 74 divisions (CI $\in [69, 79]$) as compared to 56, respectively, showcasing the versatility of the DivisionCounter method. That is, by changing the length of the microsatellite, the division range can be fine-tuned to suit the biological context.

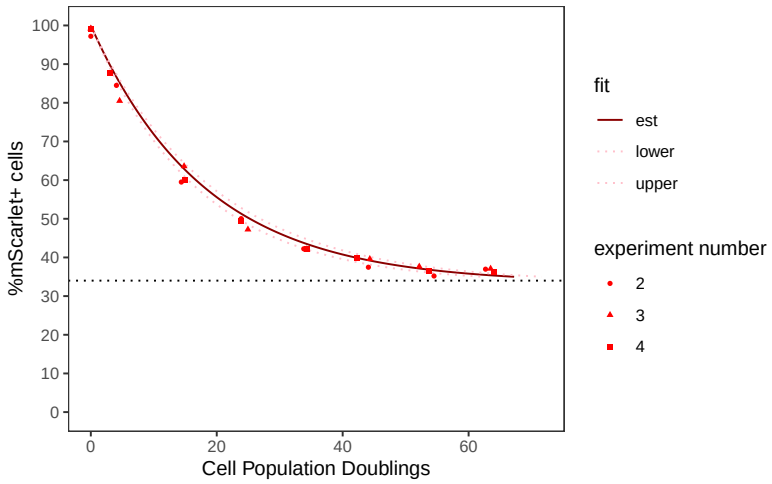
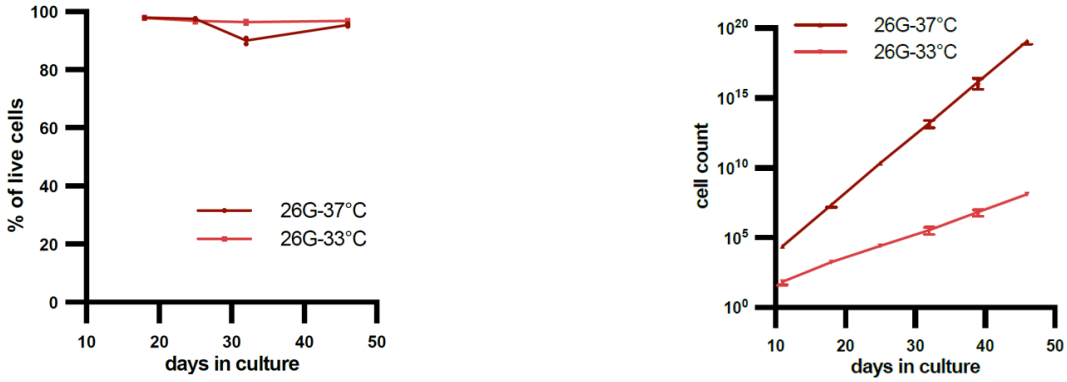


Figure 4.6: The probability of mutation for 23G-DivisionCounter.

Percentage of mScarlet-positive cells over the average decision counts measured using cell counts (population doublings) for the **23G-DivisionCounter**. Each point represents an independent acquisition, from three experiments done in triplicates. The probability of losing mScarlet-expression ρ, was fitted by a nonlinear mixed-effects model. The dark red full line shows the best fit, and the light red dot lines show the lower and upper basic bootstrap bound of 95% CI, respectively. The black line is the plateau at 35%.

To confirm that the DivisionCounter mScarlet expression was only linked to cell division, cell growth was modulated by culturing HEK293T cells infected with the 26G-DivisionCounter at 33°C to slow down cell division [86]. Figure 4.7 (a) and (b) showed cells cultured at 33°C grew slower than those at 37°C with no impact on cell mortality was confirmed, and then the percentage of mScarlet-positive cells over 49 days was accessed in *in vitro* culture.

Using the DivisionCounter formula and the nonlinear mixed-effects model to fit the probability of loss of mScarlet-expression from HEK293T cultured at 33°C, the best estimate was 0.0478 and the CI $\in [0.0470, 0.0485]$. The probability of mScarlet-expression loss at 33°C was close to the probability at 37°C ($\rho = 0.047$). Using the DivisionCounter formula and the probability of mScarlet loss from HEK293T cultured at 37°C to compute the



(a): Percenatage of cells alive (b): Cell counts

Figure 4.7: HEK293T cells mortality in 37°C and 33°C.

(a) The percentage of HEK293T cells alive over time in culture at 37°C (dark red line) and 33°C (red line). The live cells are defined as Sytox green negative cells. (b) **26G-DivisionCounter** infected HEK293T cell counts over time in culture at 37°C (dark red line) and 33°C (red line) measured using the Septor cell counter. Mean and standard deviation displayed from three experiments done in triplicates.

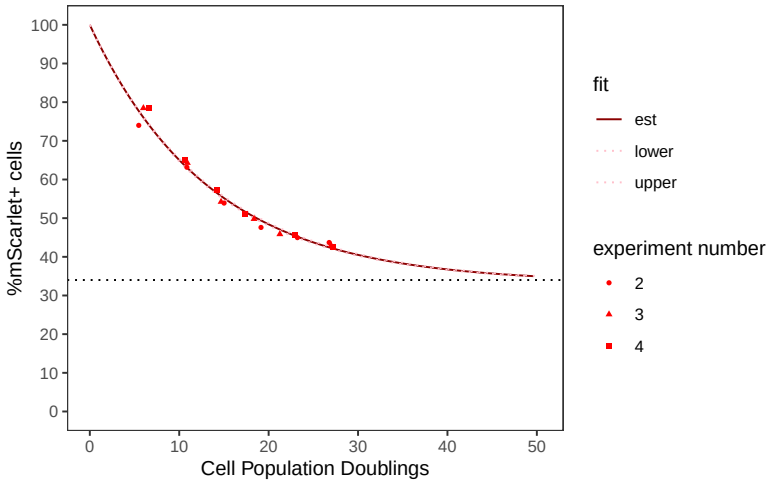


Figure 4.8: The probability of mutation for 26G-DivisionCounter in 33°C.

Percentage of mScarlet-positive cells over the average decision counts measured using cell counts (population doublings) for the **26G-DivisionCounter** in 33°C. Each point represents an independent acquisition, from three experiments done in triplicates. The probability of losing mScarlet expression ρ, was fitted by a nonlinear mixed-effects model. The dark red full line shows the best fit, and the light red dot lines show the lower and upper basic bootstrap bound of 95% CI, respectively. The black line is the plateau at 35%.

average division count at 33°C, the cell population doublings were predicted accurately with a Pearson correlation of 0.99, as shown in Figure 4.9. These results further validated that changes in the percentage of mScarlet-positive cells were exclusively linked to cell division. Importantly, the DivisionCounter can quantitatively measure the division rate.

Fitting a linear regression to the average division counts over time, cells cultured at 33°C growth slower than cells cultured at 37°C in Figure 4.10. In detail, the division rate at 37°C is 1.32 and CI $\in [1.25, 1.39]$, and the division rate at 33°C is 0.59 and CI $\in [0.58, 0.60]$. It indicated cells cultured at 33°C grew approximately 2.24 times slower than cells cultured at 37°C. These results demonstrated that the DivisionCounter effectively detects and quantifies variations in division rates, showcasing its utility for quantifying modulation of division rate across conditions by simply measuring FP expression loss.

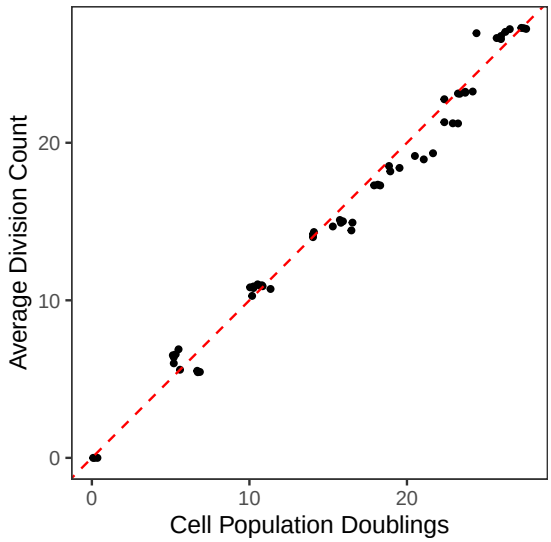


Figure 4.9: Using ρ at 37°C to predict average division count at 33°C.

The **26G-DivisionCounter** infected HEK293T cells of mScarlet-positive at 33°C. The y-axis was the average division count calculated using the percentage of mScarlet-positive cells at 33°C into the Division Counter formula with the probability of mScarlet loss at 37°C. The x-axis was the cell population doublings at 37°C. The Pearson correlation was 0.99. The red dot line showed the y=x.

To further validate the DivisionCounter method, its performance was compared to another state-of-the-art division counting method: the Cell Trace Violet (CTV) dilution dye. MEF cells were double-labelled with the **26G-DivisionCounter** construct and CTV was then cultured, counted and assessed for their mScarlet and CTV status by flow cytometry over

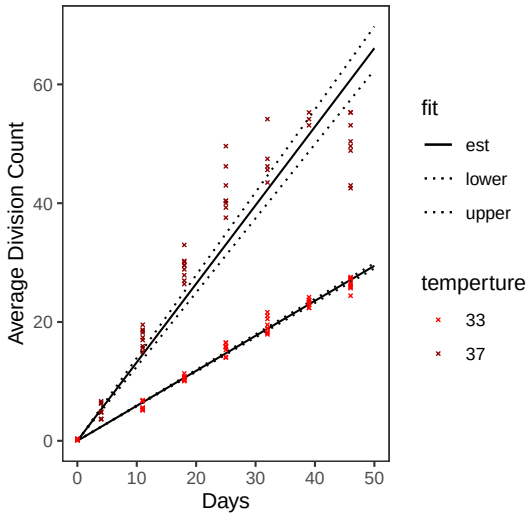


Figure 4.10: Division rates at 37°C and 33°C.

DivisionCounter average division count changes over time in HEK293T cells cultured at 37°C (dark red) and 33°C (red). The probability of loss of mScarlet-expression used in DivisionCounter formula comes from inference made on HEK293T cells cultured at 37°C. Each point represents independent acquisition from three experiments done in triplicate. The full lines represent the best-fitted linear model and the dotted lines are the confidence interval for the model.

51 days in Figure 4.11. Over the lower range of divisions that can be accurately followed with the two methods (i.e., below 10 divisions), the DivisionCounter average division count was observed to be strongly correlated with that of the CTV dye dilution with Spearman coefficient as –0.93 in Figure 4.11, which further demonstrated the reliability and accuracy of the DivisionCounter method. Notably, however, CTV reached its dilution limit after 10 divisions, directly showing that the DivisionCounter method outperformed the dye dilution state-of-the-art method.

In conclusion, the DivisionCounter method was validated and benchmarked, demonstrating its utility and accuracy to measure division counts and division rates. The DivisionCounter allows up to 74 divisions to be robustly counted, surpassing the 10-division limit of other state-of-art methods and the Hayflick limit of 40-60 divisions.

4.4.2 DivisionCounter measures B-ALL cell division rates across organs in vivo

After confirming the performance of DivisionCounter *in vitro*, this section employed the well-established Nalm6 leukaemia xenograft model in immunodeficient NSG mice to eval-

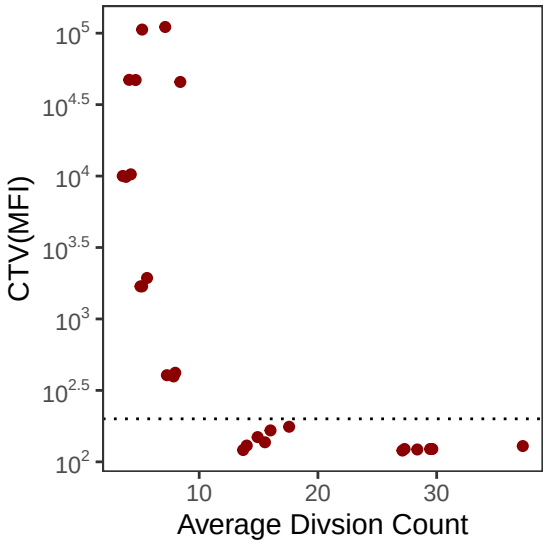


Figure 4.11: Compare CTV and the average division count predicted by DivisionCounter formula.

Median fluorescence intensity (MFI) of the cell trace division in **26G-DivisionCounter** versus the DivisionCounter average division counts computed using the **26G-DivisionCounter**. The area below the black dotted line represents the limit of detection of CTV. Each point represents an independent acquisition.

uate the performance of DivisionCounter *in vivo*. All biological experiments in this section were designed and conducted by our collaborators at Curie Institute, and all statistical analyses were performed by us.

In this tumour model development, Nalm6 B-cell acute lymphocytic leukaemia (B-ALL) cells progressively grow and spread to various organs, initially in the bone marrow (BM) and liver, and then in the spleen and lungs. While different organs exhibit distinct tumour invasion movements, it is unclear whether these different environments directly influence Nalm6 growth. To investigate the impact of the organ microenvironment on Nalm6 cell divisions, the DivisionCounter method was applied in this *in vivo* tumour model.

After injecting 2.5×10^5 Nalm6 cells infected with the 26G-DivisionCounter and 26N-DivisionCounter into NSG mice, Nalm6 tumour development was followed over time at days 7, 13 and 20 after *in vivo* injection using the DivisionCounter and standard measurements of *in vivo* cell growth (luciferase assay, KI67 staining, and cell counting), which has been shown in Figure 4.12. Luciferase measurements enabled the monitoring of cancer cell metastasis in each mouse and the average radiance of Nalm6 cells (p/s/cm^2/sr) in

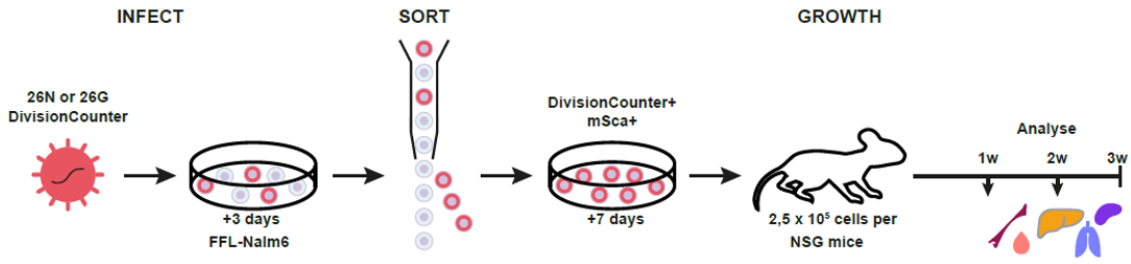


Figure 4.12: Nalm6 *in vivo* experiment.

The mTagBFP2^{+}Luciferase^{+} Nalm6 cells were infected with the 26G-DivisionCounter or 26N-DivisionCounter lentivirus at a low multiplicity of infection. After 3 days of culture, mScarlet-positive cells were sorted injected and cultured for 7 days before 2.5 x10^{5} 26G-DivisionCounter infected Nalm6 cells were intravenously injected in NSG mice. Nalm6 tumour development was then followed over time by isolating cells at day 7, 13 and 20 after *in vivo* injection from the blood, bone marrow, spleen, liver and lungs.

each mouse was obtained using In Vivo Imaging System (IVIS) [36].

Figure 4.13 showed exponential growth of Nalm6 throughout the whole body but reached saturation at day 20, preventing any further measurement of Nalm6 total body growth. The permutation test was employed to detect the null hypothesis that the average radiance of Nalm6 on the same day was independent of whether the cells were infected with 26G-DivisionCounter or 26N-DivisionCounter. Monte Carlo approximation with 500,000 test statistics was used to estimate the left-tailed p-value. The left-tailed p-values were 0.0875 and 0.3529 for day 7 and day 13, respectively, at the Bonferroni-corrected significant level ($p < 0.05/2 = 0.025$). There was no evidence to show the difference in the growth rate between 26G-DivisionCounter and 26N-DivisionCounter during the first two days. The permutation test detail is seen in Section 4.3.2.

The Luciferase measurement cannot address whether different organ microenvironments affect Nalm6 growth, illustrating its limitations. Nalm6 cell counts measured by flow cytometry with counting beads showed a higher number of cells in the BM and liver than in the lungs, blood or spleen across all time points, which has been shown in Figure 4.14. The solid dots represented the cells were infected with 26G-DivisionCounter and the empty dots meant the cells were infected with 26N-DivisionCounter. Permutation tests and Monte Carlo approximation were used and the left-tailed p-values reported in Table 4.4. There was no evidence to show the difference in the cell counts between 26G-DivisionCounter and 26N-DivisionCounter at day 7 and day 13. It also indicated the growth rate between 26G-DivisionCounter and 26N-DivisionCounter during the first two

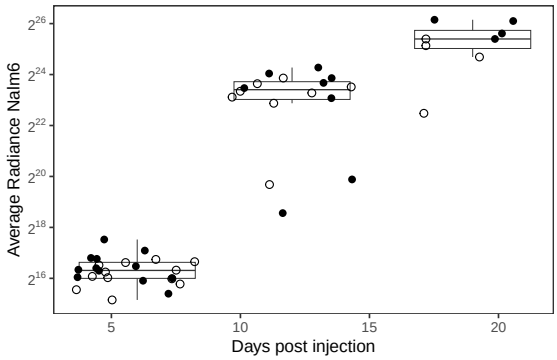


Figure 4.13: Average radiance of Luciferase measurement.

Average radiance of Luciferase measurement in the 26G-DivisionCounter (solid dots) and the 26N-DivisionCounter (empty dots) infected Nalm6 cells at day7, day13 and Day20 after *in vivo* injection.

days was the same across all organs.

In Figure 4.14, the Nalm6 cells in BM can be seen to have grown exponentially in the first two days, but the number of Nalm6 at day 20 did not increase as much as in the first two days. Nalm6 in the liver, lungs and blood grew exponentially across all time points. The number of Nalm6 cells in the spleen was much lower than that in BM and liver in the first two days, and the number of cells showed a dramatic increase on day 20. These results were consistent with previously published kinetics of tumour invasion [10].

Permutation tests and Monte Carlo approximation were also used to statistically challenge differences in the cell count between different organs on the same day in Figure 4.14. The null hypothesis was the mean of the cell count was independent of whether in different organs on the same day. Table 4.5 reported the left-tailed p-value results. From the permutation test results, there was no evidence to show that BM and liver cell counts were different on day 7 and day 13. Also, on day 7, cell counts in the lung, blood, and spleen were similar, whereas on day 13, cell counts in the blood and spleen were lower than those in the lung. These differences in cell counts indicated higher net cell growth in the BM and liver as compared to the other organs.

There was no statistical difference in cell growth between 26G-DivisionCounter and 26N-DivisionCounter Nalm6 cells based on luciferase and cell count measurements. This validates that DivisionCounter does not affect the growth properties of Nalm6 cells, demonstrating the usability and robustness of DivisionCounter for *in vivo* use.

However, it was impossible to conclude if these differences in Nalm6 cell counts were due to faster division rate, delayed or lower initial cell seeding, or differences in migration between organs.

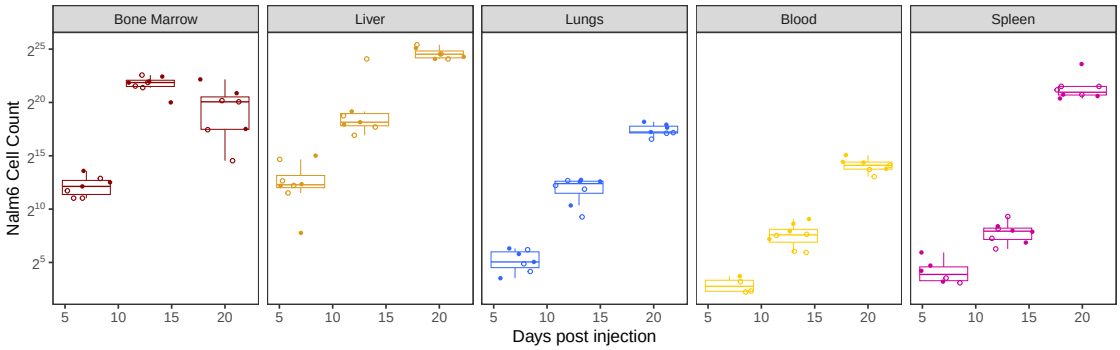


Figure 4.14: Nalm6 cell counts at day 7, 13, 20 in different organs.

Nalm6 cell counts ($CD45^{-}HLA^{+}$) over time were in bone marrow, liver, lungs, blood, and spleen respectively at day 7, day 13 and day 20. The solid dots were in 26G-DivisionCounter, and the empty dots were in 26N-DivisionCounter.

	BM	**Liver**	**Lungs**	**Blood**	**Spleen**
day 7	0.11	0.43	0.43	0.25	0.13
day 13	0.59	0.57	0.26	0.043	0.61

Table 4.4: Left-tailed p-values for cell counts in the 26G-DivisionCounter and 26N-DivisionCounter.

Since existing methods have shown their limitations, DivisionCounter was used. First, the loss of a percentage of mScarlet-positive cells of Nalm6 was obtained using *in vitro* experiments. The use of the 26G-DivisionCounter *in vitro* was validated on $mTagBFP2^{+}$ $Luciferase^{+}$ Nalm6 cells. After infection, Nalm6 cells were cultured, counted and assessed for the percentage of mScarlet-positive cells, for up to 59 days. As expected, the percentage of mScarlet-positive Nalm6 decreased with the cell population doubling and reached the 1/3 equilibrium in Figure 4.15. In contrast, in the 26N-DivisionCounter Nalm6 cells culture, the percentage of mScarlet-positive cells remained stable over time, which was in grey points. This validated the use of the DivisionCounter in Nalm6 cells. Fitting a nonlinear mixed-effects model, the probability of losing mScarlet-expression in Nalm6

Day	p-value	Liver	Lungs	Blood	Spleen
Day7	**BM**	0.2	0.0003	0.0029	0.00056
	Liver		0.00016	0.0021	0.0003
	Lungs			0.006	0.07
	Blood				0.048
Day13	**BM**	0.04	0.00009	0.00009	0.00007
	Liver		0.00013	0.00016	0.00014
	Lungs			0.0001	0.00015
	Blood				0.35

Table 4.5: The left-tailed p-values for cell counts in different organs.

cells was obtained at 0.049 and the CI $\in [0.046, 0.057]$, similar to the HEK293T cells in 26G-DivisionCounter at 37°C in Figure 4.4. The maximum number of generations tracked by DivisionCounter for Nalm6 was 54, with a CI of 45 to 57 divisions.

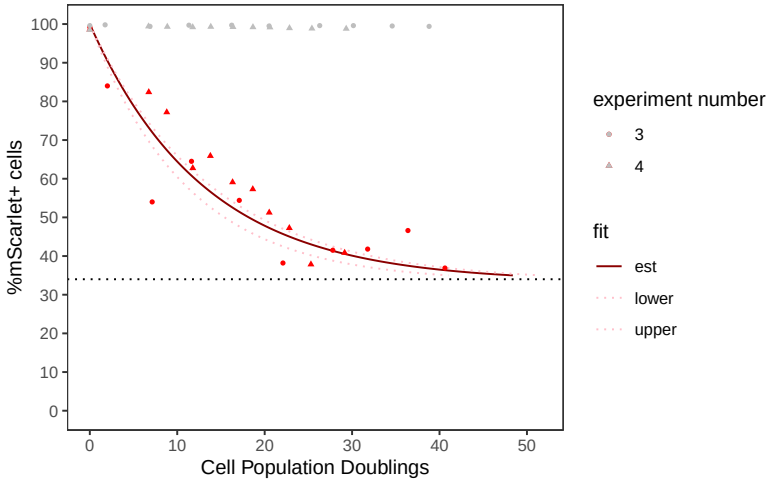


Figure 4.15: The probability of mutation of Nalm6 for 26G-DivisionCounter.

Percentage of Nalm6 mScarlet-positive cells over the number of divisions measured using cell counts (population doubling) for the 26G-DivisionCounter (red) and 26N-DivisionCounter (grey). Each point represents independent acquisition from three experiments. The probability of losing mScarlet-expression, ρ, was fitted using a nonlinear mixed-effects model and the confident interval obtained from bootstrapping.

After obtaining the best estimate of the probability of losing mScarlet-expression of Nalm6 cells, the number of cells that need to be measured was assessed. This ensured that the average division count estimate had a given error margin level with respect to the true

percentage of mScarlet-positive cells for the 26G-DivisionCounter construct. The x-axis of Figure 4.16 showed the average generation as determined by the value of the percentage of Nalm6 mScarlet-positive cells, and the y-axis was the number of cells required with a 5%, 10%, 15% and 20% margin of error on the lower tail of the estimate at 95% confidence interval. Table 4.6 showed the minimum number of cells required at different margins of error.

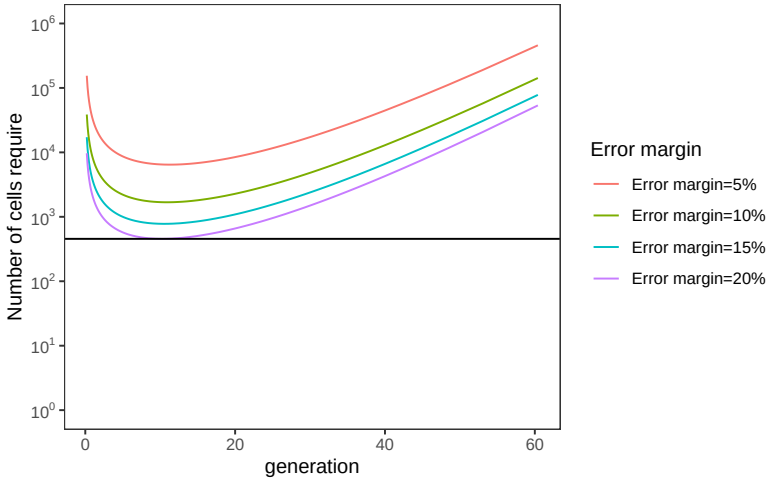


Figure 4.16: Simulation for the lower tail of the number of cells required *in vivo*.

The average division count versus the number of cells that must be acquired to ensure that the measured division count is within a given percentage of the true one. For the best estimate of the probability of losing mScarlet expression, if 455 (the black line) or more cells are measured, there is at least a 97.5% chance of being with 20% of true value.

Margin of error	**5%**	**10%**	**15%**	**20%**
Number of cells required	6450	1682	778	455

Table 4.6: Minimum number of cells required by simulation.

Next, the DivisionCounter method was tested to measure Nalm6 proliferation *in vivo*. After obtaining luciferase and cell count measurements at days 7, 13, and 20 post-injection, the percentage of mScarlet-positive cells was collected for mice in the 26G-DivisionCounter construct, excluding the number of Nalm6 cells below 455.

To evaluate the impact of the organ microenvironment on Nalm6 cell division, the *in vivo* average division counts obtained by inputting the percentage of mScarlet-positive

cells into the DivisionCounter formula were compared in the different organs. The average division counts were similar between the BM, liver and lungs at days 7 and 13 in Figure 4.17. Similarly, by day 20, when the Nalm6 had spread well into all organs, the DivisionCounter measured on average 21-26 divisions in all organs.

The permutation tests were used to statistically assess the average *in vivo* division counts obtained with DivisionCounter on the same day in Figure 4.17. The null hypothesis was that on the same day, the mean of the average division count obtained with DivisionCounter was independent of the organ. The left-tailed p-values were calculated by Monte Carlo approximation with 500,000 sampling. The test results were shown in Table 4.7, and the Bonferroni-corrected p-value threshold of $p < 0.05$ at day 7, $p < 0.05/3 = 0.0167$ at day 13, and $p < 0.05/10 = 0.005$ at day 20. There was no evidence to show the statistical difference between organs' average division count. These results indicated that the microenvironment provided by different organs did not impact the Nalm6 division. It also meant there was no evidence to show the statistical difference between organs' division rates.

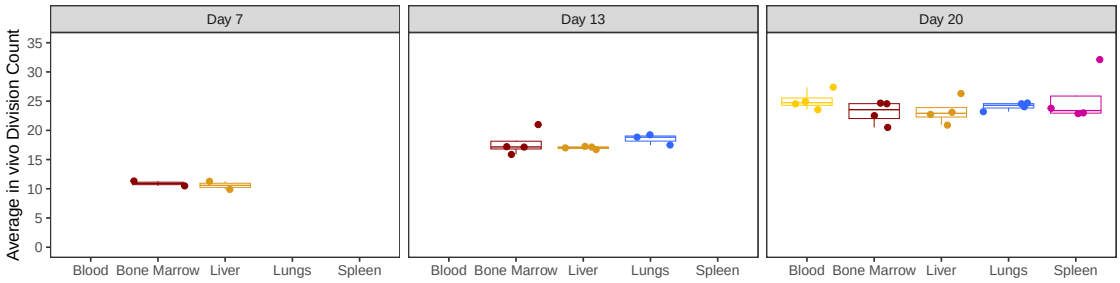


Figure 4.17: Average division count got by DivisionCounter on different days.

DivisionCounter average division count computed using the best estimate of the probability of mutation of Nalm6 at day 7, day 13 and day 20 in different organs.

Next, a simple linear regression was employed to assess the division rate in BM, liver and lungs. Here, the two day 0 *in vitro* data were added as an estimate of the initial average division count. In detail, using the probability of losing the mScarlet-expression of $BFP^+Lucipherase^+$Nalm6 cells infected with 26G-DivisionCounter *in vitro* as the parameter to fit the DivisionCounter formula, average division counts *in vivo* were predicted by the percentage of mScarlet-positive cells. The output variable of linear regression was the average division counts obtained with the DivisionCounter formula and the input variable was the time of days post-injection. A bootstrap sample with 100,000 datasets was cre-

Day	p-value	Liver	Lungs	Blood	Spleen
Day7	**BM**	0.33	-	-	-
	Liver		-	-	-
	Lungs			-	-
	Blood				-
Day13	**BM**	0.356	0.342	-	-
	Liver		0.0283	-	-
	Lungs			-	-
	Blood				-
Day20	**BM**	0.42	0.2	0.1	0.26
	Liver		0.26	0.11	0.27
	Lungs			0.19	0.49
	Blood				0.49

Table 4.7: Left-tailed p-values for average division in different organs.

ated to get 95% confidence interval by sampling with experiment number and technical repeat from the *in vitro* dataset, and time from the *in vivo* dataset.

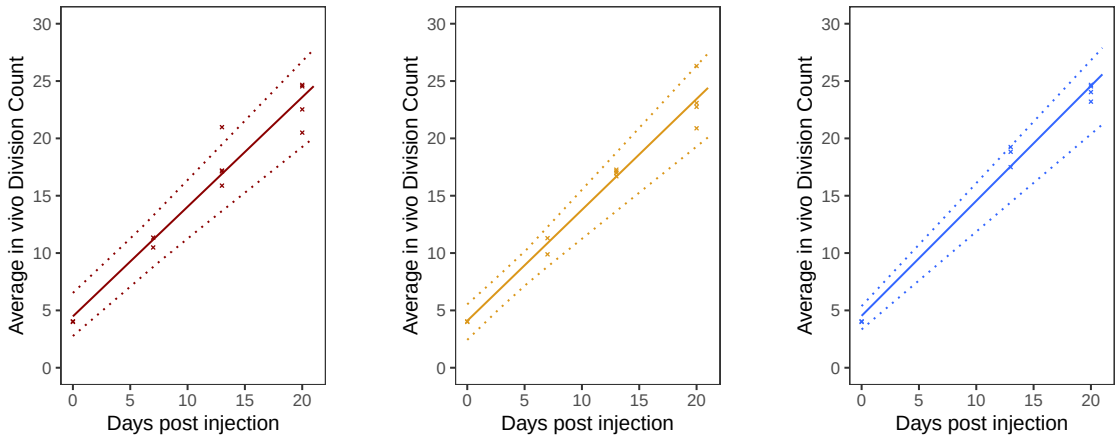
Figure 4.18 revealed that in all organs and at all time points, Nalm6 cells had divided more than 10 times. Using the DivisionCounter, Nalm6 cells were determined to have divided around 11, 18, and 23 times at days 7, 13 and 20 post-injection, respectively, allowing for the first time to measure the division counts of tumour *in vivo*. These values could not have been measured by any existing state-of-the-art method. Overall, the DivisionCounter method demonstrated high sensitivity for *in vivo* counting of divisions.

	Bone marrow		**Liver**		**Lungs**	
	Intercept	**Coefficent**	**Intercept**	**Coefficent**	**Intercept**	**Coefficent**
Lower	3.8	0.80	3.3	0.80	3.9	0.85
Estimated	4.5	0.96	4.1	0.97	4.5	1.00
Upper	4.2	1.05	3.4	1.09	5.1	1.07

Table 4.8: Fitted Nalm6 division rate in each organ.

Table 4.8 showed the fitted Nalm6 division rate in each organ. The DivisionCounter method revealed that Nalm6 cells in BM, liver and lungs all divided at a rate of around one division per day. Table 4.7 showed no statistical difference between these three organs. Thus, the DivisionCounter demonstrated that the *in vivo* division rates of Nalm6 leukaemia tumour cells in NSG mice were not modulated by the organ-specific microenvironment.

The DivisionCounter results with cell counts were integrated to gain further biological



(a): Bone marrow (b): Liver (c): Lungs

Figure 4.18: Fitted Nalm6 division rate in each organ.

DivisionCounter average division counts were computed using the best estimate of the probability of mutation of Nalm6 over time in the bone marrow, liver and lungs. Each dot was a mouse. A linear regression was fitted to the average division count over time, and the obtained division rate with confidence interval was computed using bootstrapping.

insight into the kinetics of tumour invasion in different organs. The cell counts measured by flow cytometry with counting beads were the output data. The data were fitted by the exponential model with the base of 2, and the exponential model with two parameters initial value and growth rate. Bootstrap samples were created by sampling with time to get 95% confidence intervals. Figure 4.19 and Table 4.9 showed the fitted results.

The day 20 BM cell counts were predicted by day 7 and day 13 cell counts employing the best fit exponential model for BM and spleen. The permutation test was used to challenge the cell counts at day 20 and the predicted cell counts at day 20. The null hypothesis for the permutation test for BM was that the mean number of cell counts at day 20 was independent of whether the cell counts were predicted or not. The p-value for the left-tailed test was estimated to be 0.00001 using a Monte Carlo approximation with 500,000 test statistics. It means the cell counts in the BM were significantly lower than expected by exponential growth on day 20. From Figure 4.19, between days 13 and 20, Nalm6 cell counts stopped increasing in the BM but continued to increase in the other organs.

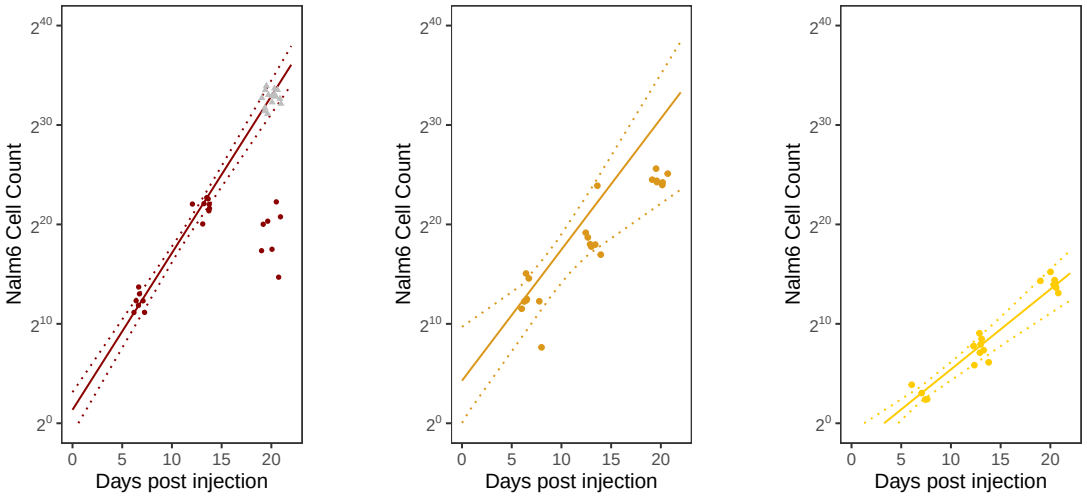
The plateauing of Nalm6 cell counts in the BM could be due to a cessation of cell division or increased cell death, but cell counts alone cannot resolve those assumptions. The DivisionCounter, however, revealed that Nalm6 cells in the BM continued to divide at the same rate between days 13 and 20, with no change in the percentage of dead cells observed at day 13 and 20, demonstrating that the lower number of Nalm6 cells in the BM was not due to change in division or death.

An alternative assumption is that the tumour cells lacked space in the BM and egressed to compensate for their growth, as observed in an acute myeloid leukaemia mouse model. In support of this assumption, the day 20 spleen cell counts were predicted by day 7 and day 13 cell counts with the best fit exponential model for the spleen. The null hypothesis for the permutation test for the spleen was that the mean number of cell counts at day 20 was independent of whether the cell counts were predicted or not. Monte Carlo approximation with 500,000 test statistics was used to estimate the p-value for the right-tailed test, and the p-value was 0.000002. It meant a significant increase in Nalm6 cell counts in the spleen at day 20 was observed, higher than expected due to exponential growth, along with the presence of Nalm6 in the blood, both suggesting increased circulation of Nalm6 cells between organs.

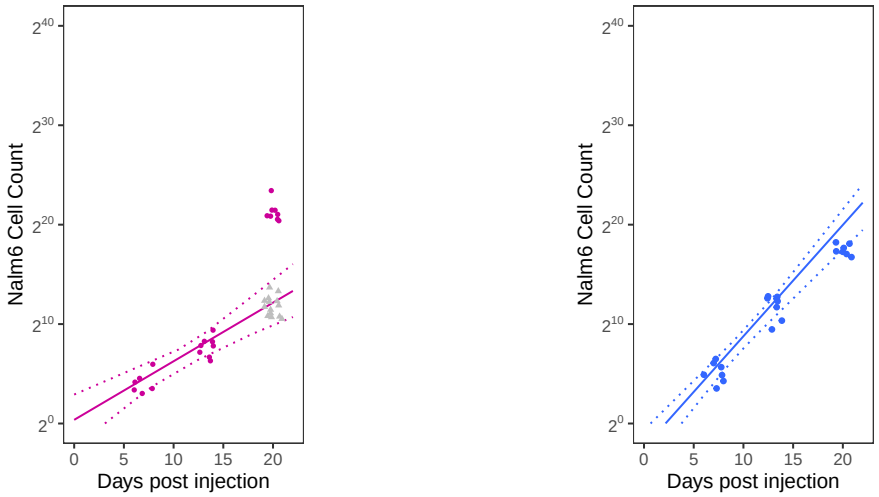
Therefore, the stable Nalm6 cell counts in the BM would be well explained by a BM space saturation coupled with the migration of Nalm6 cells into the circulation and spleen [4]. Additionally, lower Nalm6 cell counts in the lungs were observed at days 7 and 13 as compared to BM and liver in Figure 4.14. The DivisionCounter method showed that Nalm6 cells had the same average division counts at days 7, 13 and 20 in lungs, BM and liver, with no significant difference in division rates between the lungs and liver at day 20 in Figure 4.17, the differences in division rate were not responsible for the lower cell count in the lungs. It is more likely that the lower Nalm6 cell counts in the lung were due to a lower number of Nalm6 cells seeding the lungs, or to a later tumour seeding in the lungs from Nalm6 cells migrating from the BM or liver via the circulation. Therefore, the DivisionCounter is more informative than cell counting alone, and using the DivisionCounter in conjunction with cell counting significantly enhanced our understanding of the kinetics and spatial dynamics of Nalm6 tumour development *in vivo* across multiple organs.

4.4.3 DivisionCounter measures the in vivo tumour division rate during CAR-T cell therapy, allowing the CAR-T cell in vivo killing rate to be inferred

In this section, the utility of the DivisionCounter in deconvolving the effect of cell division and cell death in the context of targeted tumour-killing cell therapies was investigated



(a): Bone Marrow (b): Liver (c): Blood



(d): Spleen (e): Lungs

Figure 4.19: Fitted net growth in each organ for day 7 and 13 based on cell counts.

Nalm6 cell counts ($CD45^{-}HLA^{+}$) were over time in bone marrow, liver, blood, spleen and lungs. An exponential regression was fitted to the cell counts with the obtained average net cell growth rate ϖ value indicated for each organ with a confidence interval computed using bootstrapping. The grey dots are the points predicted using the fitted exponential growth rate for the bone marrow and the spleen. Each dot is a mouse, and the 26G-DivisionCounter and 26N-DivisionCounter pooled.

	Bone marrow		Liver		Blood		Spleen		Lungs	
	ϱ_{t_0}	ϖ	ϱ_{t_0}	ϖ	ϱ_{t_0}	ϖ	ϱ_{t_0}	ϖ	ϱ_{t_0}	ϖ
Lower	4.5	1.45	509	0.69	0.39	0.65	5.3	0.41	0.57	0.98
Estimated	2.5	1.58	19	1.32	0.16	0.81	1.3	0.59	0.19	1.12
Upper	0.78	1.72	3.7	1.65	0.05	0.96	0.2	0.79	0.06	1.27

Table 4.9: Fitted net growth in each organ for day 7 and 13 based on cell counts.

The ϱ_{t_0} was the initial value of exponential regression, ϖ was the growth rate, and the base of the exponential function was 2.

with experiments conducted by our collaborators at Curie Institute and data analyses performed by us.

The DivisionCounter approach was employed to explore how targeted killer cell therapy affects Nalm6 B-ALL cell proliferation *in vivo*. Adoptive cell therapies with T cells expressing chimeric antigenic receptor (CAR) have efficient clinical results against B-ALL cells [21, 22]. CAR-T cells efficiently kill tumour cells either directly or via the microenvironment [24, 30, 84, 155]. However, it is unknown if they directly impact tumour cell division. To evaluate this, the DivisionCounter was used to quantify the average division counts over time in the presence or absence of CAR-T cell therapy.

Seven days after injecting 2.5×10^5 26G-DivisionCounter Nalm6 cells into NSG mice, then 2×10^6 T cells transduced with the 4-1BB-based CD19 CAR construct or mock construct were transferred into separate mice, which has been shown in Figure 4.20.

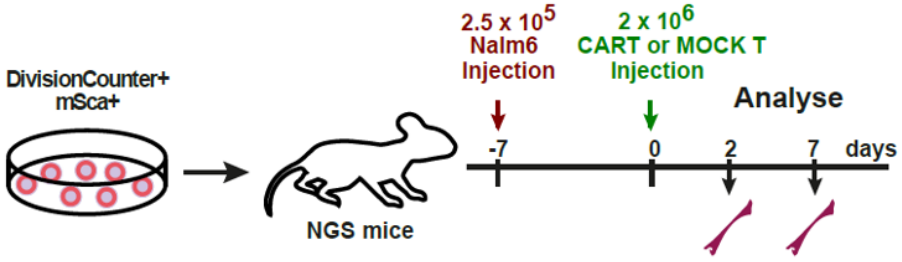


Figure 4.20: CAR-T *in vivo* experiment.

The mTagBFP2$^+$Luciferase$^+$ Nalm6 cells were infected with the 26D-DivisionCounter lentivirus at a low multiplicity of infection. After three days of culture, mScarlet-positive cells were sorted injected and cultured for seven days. The 2.5×10^5 26G-DivisionCounter infected Nalm6 cells were injected into NSG mice. Seven days later 2×10^6 4-1BB CAR-T or MOCK-T were injected. Tumour development was followed over time by isolating cells from the bone marrow at day 2, 7 and 13 post T cell injection.

Luciferase measurements were used to measure average Nalm6 radiance in mice at day -1, day 1, day 6, and day 13. Figure 4.21 showed the average Nalm6 radiance in mice injected with CAR-T and MOCK-T. From day -1 to day 6, the average Nalm6 radiance increased exponentially. A lower Nalm6 tumour burden at day 13 days post-injection with T-cells, in CAR-T cell-treated mice as compared to the MOCK-T cell-treated mice.

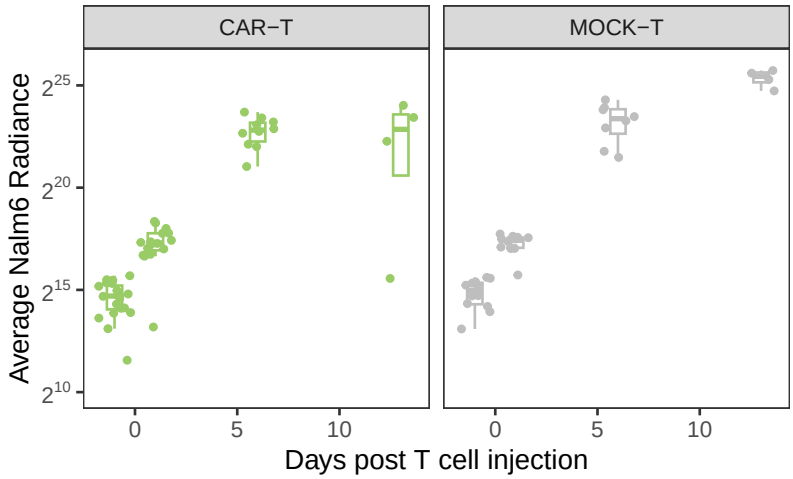


Figure 4.21: Average radiance of Luciferase measurement.

Average radiance of Luciferase measurement mice injected with CAR-T (green) or MOCK-T (grey) at day -1, day 1, day 6 and day 13 post T cell injection. The boxplots showed the distribution of average radiance of each day, and each dot represented a mouse.

The permutation test and Bonferroni correction were used to statistically assess the average Nalm6 radiance on the same day in Figure 4.21. The null hypothesis was that on the same day, the mean of the average radiance was independent of the therapy. The left-tailed p-values were calculated by Monte Carlo approximation with 500,000 samplings, and the thresholds for the significance of Bonferroni-corrected p-values was $p < 0.05/4 = 0.0125$. The test results are shown in Table 4.10. There was no evidence to show the statistical difference on the average radiance between CAR-T and MOCK-T treatment. However, the p-value on day 13 was 0.014. Although the p-value was greater than the significant value, the difference in average radiance at day 13 was still shown.

Nalm6 cell counts were collected by flow cytometry with counting beads at day 2 and day 7, which has been shown in Figure 4.22. This reduced tumour burden could already be

	Day -1	Day 1	Day 6	Day 13
p-value	0.19	0.66	0.08	0.014

Table 4.10: Left-tailed p-values for average radiance in different treatments.

observed in the bone marrow at 7 days post-injection, whereby BM Nalm6 cell counts in the CAR-T cell-treated group were lower than in the mock-T cells group at day 7.

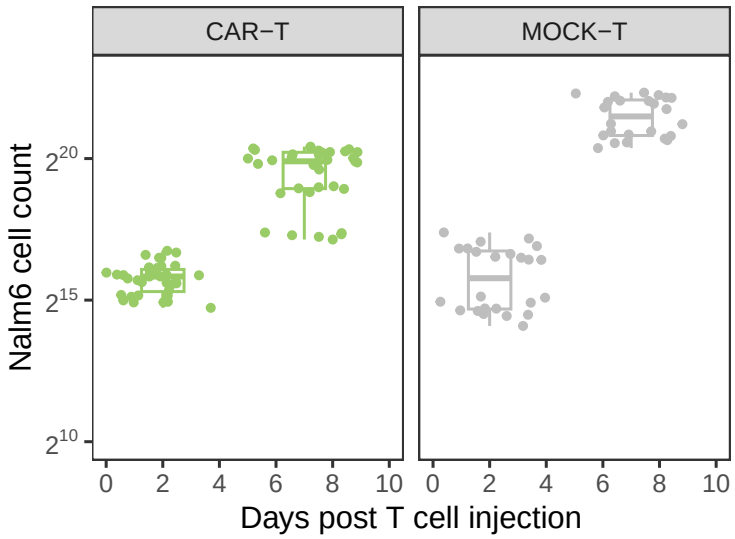


Figure 4.22: Nalm6 cell counts on CAR-T and MOCK-T.

Absolute Nalm6 cell counts ($CD45^{-}HLA^{+}$) over time in the bone marrow of mice injected with CAR-T (green) or MOCK-T (grey). Each dot was one replicate, ten replicates per mouse, six mice per time points for the CAR-T condition and four per time points for the MOCK-T condition. The boxplots showed the distribution of Nalm6 cell counts each day under different treatments.

Permutation tests were also used to assess differences in the Nalm6 cell counts in the BM of mice injected with CAR-T and MOCK-T on the same day in Figure 4.22. The null hypothesis was the mean of the cell counts is independent of whether in different therapies on the same day. Monte Carlo approximation with 500,000 test statistics was applied to estimate the left-tailed p-value. The threshold for significance of the Bonferroni-corrected p-value was $0.05/2 = 0.025$. The test results were shown in Table 4.11.

From the test results in Table 4.11, there was no evidence that the number of cell counts on day 2 was different under different therapies. On day 7, the bone marrow Nalm6

	Day 2	Day 7
p-value	0.06	0.000002

Table 4.11: Left-tailed p-values for cell counts in different therapies.

cell counts in the CAR-T cell treatment group were significantly lower than those in the MOCK-T group. It also meant the net growth in the CAR-T cell treatment group was lower than that in the MOCK-T group.

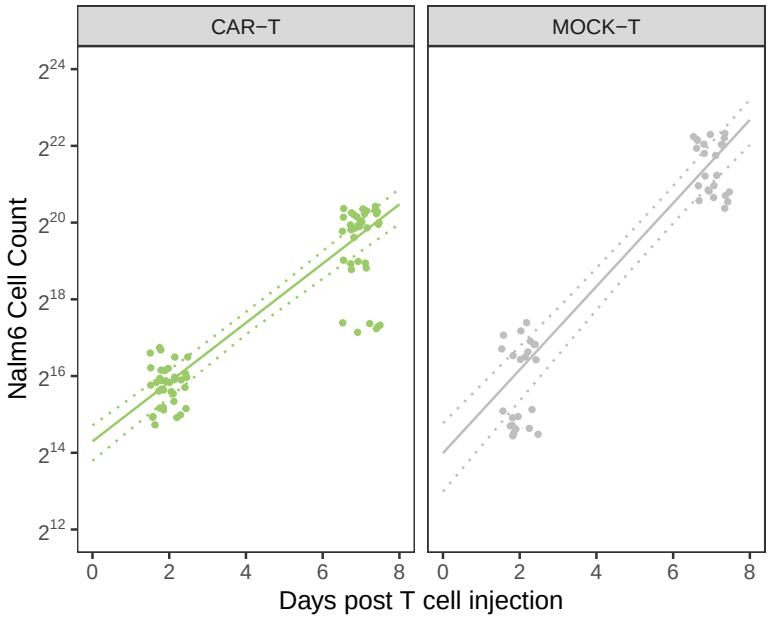


Figure 4.23: Fitted net growth in CAR-T and MOCK-T in cell counts.

Absolute Nalm6 cell counts ($CD45^{-}HLA^{+}$) over time in the bone marrow of mice injected with CAR-T (green) or MOCK-T (grey). An exponential regression was fitted to the cell counts with the obtained average net growth rate value indicated for each condition with 95% confidence interval computed using bootstrapping.

Fitting an exponential growth model onto cell counts data, and bootstrap samples were created by sampling with time to get 95% confidence interval. The model results were shown in Figure 4.23 and Table 4.12. A net growth rate of the Nalm6 cell counts was found as 0.77 cells per day in the bone marrow of the CAR-T cell treated group, which is lower than the 1.09 cell per day growth rate of the mock-T cell group. Thus, CAR-T cells were clearly killing Nalm6 and inducing reduced Nalm6 cell counts over time, yet it was still unclear whether CAR-T cells affected Nalm6 divisions.

	CAR-T		MOCK-T	
	ϱ_{t_0}	ϖ	ϱ_{t_0}	ϖ
Lower	23473	0.71	20384	0.998
Estimate	20133	0.77	16304	1.086
Upper	16222	0.83	10338	1.183

Table 4.12: Fitted net growth in CAR-T and MOCK-T in cell counts.

The ϱ_{t_0} was the initial value of exponential regression, ϖ was the growth rate, and the base of the exponential function was 2.

The DivisionCounter method assessed whether the reduced Nalm6 net cell growth in the CAR-T cell treated group could be attributed solely to CAR-T cell killing of Nalm6 or if changes in the Nalm6 division rate occurred *in vivo*. First, by inputting the percentage of mScarlet-positive cells observed in the BM over time into the DivisionCounter formula, the average division counts at 2 days post-injection and 7 days post-injection were obtained in Figure 4.24.

Similarly, permutation tests and Monte Carlo simulations were utilised to analyse the difference in Nalm6 average division counts with the DivisionCounter formula in the bone marrow of mice injected with either CAR-T or MOCK-T cells on the same day, as illustrated in Figure 4.24. The null hypothesis was the mean Nalm6 average division counts do not depend on the type of therapies, CAR-T or MOCK-T, on the same day. To evaluate the left-tailed p-value, Monte Carlo approximation was performed with 500,000 iterations, followed by a Bonferroni correction for multiple comparisons at the Bonferroni-corrected significance level ($p < 0.05/2 = 0.025$).

Table 4.13 revealed no statistical differences in Nalm6 division counts between the CAR-T cell and MOCK-T cell treated mice both on day 2 and day 7. It also demonstrated that the Nalm6 division rate was not affected by CAR-T cell killing.

	Day 2	Day 7
p-value	0.18	0.58

Table 4.13: Left-tailed p-values for average division count in different therapies.

A simple linear regression was employed to fit Nalm6 division rate in CAR-T and MOCK-T. Here, two *in vitro* data were included, which were collected at day -8. A bootstrap sample with 100,000 datasets was created to get 95% CI. Nalm6 cells divided at one

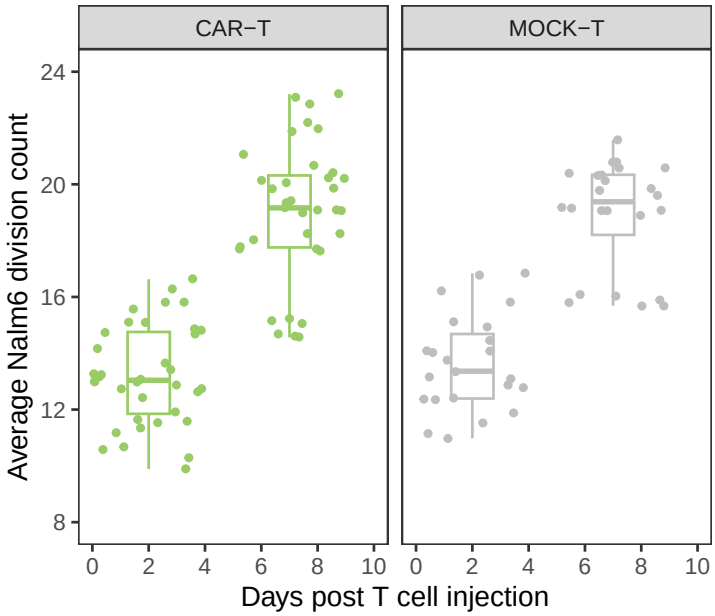


Figure 4.24: Average division count got by DivisionCounter on CAR-T and MOCK-T.

Average division count of 26G-DivisionCounter Nalm6 cell computed using the DivisonCounter formula based on the percentage of mScarlet-positive cell analysed by flow cytometry at each time point in the bone marrow of mice injected with CAR-T (green) and MOCK-T (grey). Each dot is an experimental data point. Sic mice per time points for the CAR-T condition and four per time points for the MOCK-T condition.

division per day in Figure 4.25 and Table 4.14, comparable to the rate without CAR-T injection in Figure 4.18 and Table 4.8. This finding was further supported by the fact that the Nalm6 division rate was not affected by the presence or action of CAR-T cells.

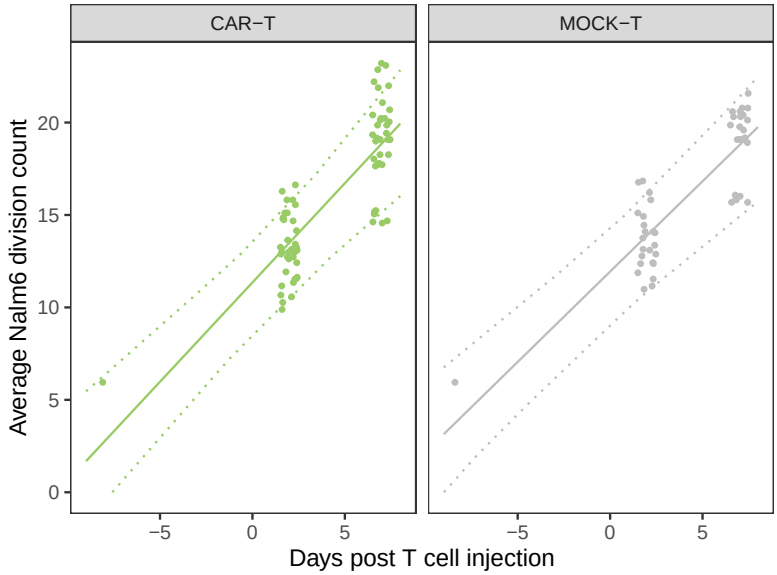


Figure 4.25: Fitted Nalm6 division rate in CAR-T and MOCK-T.

Fitted division rate of 26G-DivisionCounter Nalm6 cells over time based on the average division count in the bone marrow of mice injected with CAR-T (green) or MOCK-T (grey). The division rate is fitted with a confidence interval using bootstrapping. The lines represent the best estimate and confidence interval for the expected division counts using the fitted division rate.

	CAR-T		**MOCK-T**	
	Intercept	**Ceofficent**	**Intercept**	**Ceofficent**
Lower	9.3	0.87	12.6	0.79
Estimate	11.3	1.07	11.9	0.98
Upper	11	1.24	12	1.12

Table 4.14: Fitted Nalm6 division rate in CAR-T and MOCK-T.

Lastly, combining the DivisionCounter inferred division rate with the cell count inferred net cell growth rate, the Nalm6 cell death rate induced by CAR-T cells was estimated as 0.31 cells per day and CI $\in [0.21, 0.43]$ in Figure 4.26. This low value could be due to the sub-optimal number or potency of the CAR-T cells, demonstrating the power of the method to evaluate the factors influencing tumour growth during treatment.

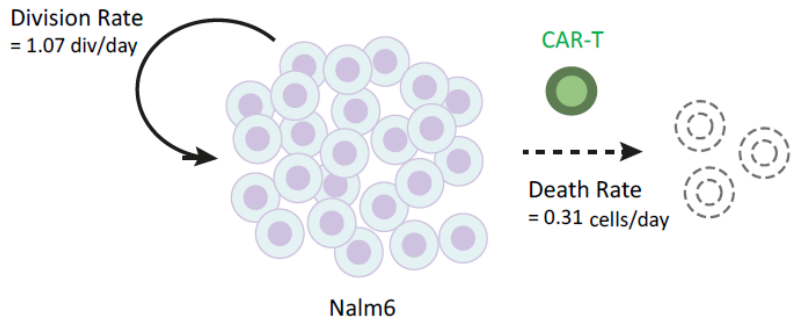


Figure 4.26: Schematic of the Nalm6 division rate and CAR-T killing rate.

In conclusion, the DivisionCounter is a robust method for quantifying cell division *in vivo* in animal models, in a cell population subject to complex cell population dynamics that include death and migration. The DivisionCounter is the first method capable of quantifying cell division and death rates *in vivo* over large division ranges, enabling the user to quantify the contributions of division and other factors, such as death and migration, to cell growth. The DivisionCounter's utility has been showcased to decipher the kinetics of tumour growth and the effect of tumour therapies.

4.5 Discussion

This study developed the DivisionCounter method, a robust, accurate and versatile cell population division counter that can quantify cell divisions and division rates over time and in different conditions. My contribution to this work was in the development of the quantitative methods necessary to interrogate the experimental data provided by collaborators in Dr. Perié's Lab. The DivisionCounter method comprises a genetic construct and a mathematical formula, and it outperformed start-of-the-art methods by counting between 56 and 74 divisions for its 26G- and 23G-versions, respectively, reaching the Hayflick limit of 40-60 divisions and the estimated division number of most cells [63]. Indeed, 44 divisions without death would produce all 10^{13} cells in a human body [28, 132].

The chief assets of the DivisionCounter include i) a quantitative framework comprising the DivisionCounter formula and statistical testing to estimate confidence interval and division rate over time; ii) a well-defined working range and accuracy. The DivisionCounter counts divisions until it reaches the equilibrium between the three states of its genetic construct and, after this point, still provides a lower bound of division count, i.e. that cells have undergone at least this division count to reach the 1/3 equilibrium; iii) a tuneable working range of division counting by changing the motif or length of the mi-

crosatellite in the DivisionCounter construct; and iv) an easy readout using fluorescence by flow cytometry and microscopy with robust values for as little as 455 cells analysed.

The utility of the DivisionCounter method for quantifying division counts and division rates *in vivo* was demonstrated, consequently allowing death rates to be estimated *in vivo*. Changes in net cell growth rate due to microenvironment, treatments or ageing have been traditionally assessed using cell counting, division tracker dyes, snapshot measurement of the cell cycle and death markers. In contrast to these methods, the DivisionCounter allows the *in vivo* division rate to be deconvoluted from the death and migration rates. This offers the potential to quantify systemic effects on growth rate with applications in CAR-T cell development, as well as drug and genetic screenings to evaluate drugs or gene knock-down effects on cell division or death.

The DivisionCounter method offers a broad applicability. For instance, it could be reengineered and expanded to non-mammalian cells, or it could be used with single-cell omic readouts to compute the average division count within cluster of cells. Additionally, more than one DivisionCounter construct could be used, each with a distinct microsatellite and FP, to provide multiple ranges of division count resolution. For cells that are difficult to transduce using lentiviruses, it would be possible to adapt the DivisionCounter method to include a genetic mouse model, similar to the ones used in mosaicism research [78]. In this chapter, the mathematical derivation of a new relationship has been described that enables the estimation of the average number of divisions a population has undergone. Biological experiments have been described that our collaborators performed to realize the method, as well as the statistical framework that we developed along with the resulting interpretation. Overall, the DivisionCounter method provides the potential to study the role of cell division in healthy and pathological tissue development *in vitro* and *in vivo*.

In this chapter, the mathematical derivation of a new relationship was described that enables the estimation of the average number of divisions a population has undergone. All involved experiments that our collaborators performed were described to realise the novel method. The statistical framework we developed was reported, which included the method to estimate the key parameter of DivisionCounter, challenge the biological hypothesis, evaluate the growth rate and division rate between different conditions, approximate the killing rate of the CAR-T therapy, and access the confidence interval of all the statistics. The statistical framework not only validates the DivisionCounter method but also demonstrates how the statistical framework can be applied to future studies of temporal generational distribution in cells, making it a useful tool for broader generational research applications.

5
Conclusion

This thesis has presented the development and application of novel statistical frameworks to address complex biological questions related to spatio-temporal cell distribution analysis. By integrating advanced statistical methodologies with cutting-edge biological techniques, this work provides new insights into cellular behaviours within the haematopoietic system. The three main studies presented in Chapter 2, Chapter 3, and Chapter 4 build upon one another, forming a cohesive narrative that demonstrates the power of statistical analysis in understanding dynamic cellular behaviours.

In Chapter 2, a spatial statistical pipeline was designed to analyse the spatial relationships between various cell types within bone marrow at a fixed time. This method provided a quantitative framework for investigating how cellular arrangements impact haematopoiesis. By going beyond the limitations of traditional random simulations and focusing on direct spatial comparisons, this pipeline yielded new insights into the interactions between AML cells, MGKs, and T cells, highlighting both known and unexpected patterns of cellular behaviour. The approach offers significant potential for further unravelling the complexities of multicellular environments and their role in disease progression.

Chapter 3 built upon this spatial analysis by incorporating a temporal dimension, introducing a spatio-temporal statistical framework to analyse high-dimensional *in vivo* microscopy data. This study focused on HSCs dynamics during early engraftment, using a data-cleaning pipeline and advanced machine-learning methods to uncover key aspects of HSCs movement and differentiation. The results challenged long-standing views on the role of the niche in cell maintenance, revealing a direct link between cell movement and differentiation status. This framework offers a foundation for future studies on how spatial proximity to niche components, such as nestin-GFP+ cells, influences cell fate decisions in dynamic environments.

Chapter 4 shifted the focus from short-term spatio-temporal analysis to long-term temporal modelling, extending the statistical analysis beyond cell interactions to the entire lifespan of normal cells. While Chapter 2 and Chapter 3 explored the spatial and short-term spatio-temporal distribution of cells, Chapter 4 introduced the DivisionCounter method to address the challenge of measuring cell divisions over long periods *in vivo*. By combining a biological tool with a novel mathematical and statistical framework, the method estimated cell divisions over 70 generations. The DivisionCounter method could also distinguish between cell division and cell death, making it a powerful tool for studying tumour growth, treatment effects, and other processes related to cell behaviour. Its ability to be applied across various biological analyses makes the DivisionCounter a valuable tool for advancing research in biological studies of long-term cellular behaviour. This chapter focused on temporal analysis, broadening the statistical framework to track how cells change over time. It complements the spatio-temporal analyses in the earlier chapters.

Together, the statistical work presented in this thesis highlights the importance of integrating spatio-temporal statistical methodologies with biological data to generate meaningful inferences about cellular behaviour. The framework developed is not only applicable to the specific contexts explored in these works, but also provides decision support for a wide range of future studies by biologists and clinical applications.

Moving forward, the methodologies developed in this thesis could be expanded beyond the specific biological contexts explored here. The spatial statistical pipeline from Chapter 2 can be applied to other cell types and tissues where cell interactions play a critical role, such as the tumour microenvironments or immune cell microenvironments. The spatio-temporal statistical framework from Chapter 3 could be further refined to investigate other dynamic cellular processes, including stem cell transplantation and tissue regeneration. Finally, the DivisionCounter method from Chapter 4 has broad applicability in studying long-term cell proliferation across different disease models, including cancer progression and response to therapy.

The work presented in this thesis highlights the importance of integrating statistical methodologies with biological data to extract meaningful insights into cellular behaviour. By developing practical, reproducible and generalizable statistical analysis methods, this research not only advances our understanding of haematopoiesis but also provides a foundation for future studies across a wide range of biological and clinical applications. This work demonstrates how statistical analysis can drive discoveries in biological research, opening new possibilities in cell biology, disease modelling, and therapy development.

A

Experiments performed by collaborators

This appendix introduces the data sources of this thesis, mainly describing the biological experimental system and data extraction process, which were done by collaborators. This section is included to ensure the thesis is self-contained, but the author takes no credit for the content.

Section A.1 describes the experiments and data extraction in Chapter 2, and these experiments were designed, produced and performed by George Adams, Floriane S. Tissot and Cristina Lo Celso. In Section A.2, the experiments are reported in Chapter 3 performed by Reema Khorshed, Sara Gonzalez-Anton and Cristina Lo Celso. Section A.3 provides the biological experiments from Chapter 4, and these experiments were designed, produced and performed by Lucie S. P. Hustin, Cecile Conrad, Jaime Fuentealba, Silvia Menegatti, Seva Shneer, Aude Battistella, Fanny Tabarin, Tom Weber, Sebastian Amigorena, Ken R. Duffy, and Leïla Perié.

A.1 PACESS experiments

A.1.1 Experimental model

Mice

All animal work was in accordance with the animal ethics committee (AWERB) at Imperial College London, UK and UK Home Office regulation (ASPA, 1986). All mice were bred and housed at Imperial College London or Sir Francis Crick Institute. C57BL/6 WT were purchased from Charles River (United Kingdom). vWF. Male and female mice $>$ 6weeks old were used. Animals were housed in Tecniplast mouse greenline cages with appropriate bedding and enrichment. The temperature, humidity and light cycles were kept within the UK Home Office code of practice, with standard diet and water ad libitum, the temperature between 20 and 24°C, the room humidity at 45–65% and a 12-hours light

/12-hours dark cycle with a 30-min dawn and dusk period to provide a gradual change.

AML experimental model

Murine AML cells were generated as described in [40]. Briefly, granulocyte/monocyte progenitors (GMPs) were purified from mT/mG mice, transduced with pMSCV-MLL-AF9-GFP retroviruses as described by [79] and transplanted into sub lethally irradiated mice (two doses of 3.3 Gy, at least 3 hours apart). At 8+ weeks post-transplantation, recipient mice developed leukemia characterised by multi-organ infiltration. tdTomato positive cells were harvested from BM and spleen and cells from each primary recipient were labelled as a separate batch and cryopreserved. Primary cells from were thawed, suspended in PBS and 100,000 viable cells were injected i.v. into secondary, non-conditioned recipient mice.

Intravital microscopy

Intravital microscopy was performed using a Zeiss LSM 980 upright confocal microscope equipped with 5 Argon lasers (405, 488, 561,594 and 639nm), and an Insight (Newport Spectraphysics) 2-photon laser with two excitation lines of which one is fixed and one tunable (1045nm and 680-1300nm respectively), 6 non-descanned external detectors including 2 nose-piece detectors (GaASP). Images were acquired using a 20x, 1.0N.A., water immersion lens with 1.4mm working distance. Live imaging of the calvarium bone marrow was performed as previously described in [62, 145]. Dextran FITC and tdTomato were excited using the 488nm and 561nm laser, respectively and detected with internal detectors.

Immunofluorescence staining on thick bone sections

Tibias, femurs and calvarium were harvested and fixed for different period of time in 4% formaldehyde at 4°C. Bones were decalcified for 10 to 15 days in 10% EDTA and tibias and femur were embedded in 4% low EEO agarose (Sigma, A0169) and cut using a Leica T1000 Vibratome at depths of $250\mu m$. All the following steps were performed under agitation at room temperature. Sections or whole calvarium were incubated in 20% CUBIC-1 reagent (urea (25 wt% final concentration), Quadrol (25 wt% final concentration), Triton X-100 (15 wt% final concentration) in dH2O) for 24 hours [140], rinse in TBS, unspecific antigen binding site were block using blocking buffer solution (TBS 0.1% Triton, 10% DMSO, 5% normal donkey serum) over nigh. The bones were then incubated with primary antibodies (CD8 clone 2.43 Invitrogen), diluted in blocking buffer for 48 hours. After several consecutive incubations in washing buffer (TBS 0.1% Triton), bones were incubated with secondary antibodies for 48 hours, followed by washes. Nuclei

were counterstained using DAPI incubation overnight. Finally, the bones were mounted in the clearing solution [85] (Methlyacetamide 40% (sigma) diluted in TBS, 1.455g histodenz (sigma) per 1ml 40% Methlyacetamide, 4% DABCO (sigma)) using silicon isolator (ThermoFischer P18175) on Superfrost Plus^{TM} slides and imaged at least 24 hours after, once the tissue is cleared.

Image acquisition

Image acquisition was performed using a Zeiss LSM 980 upright confocal microscope equipped with 5 Argon lasers (405, 488, 561,594 and 639nm), and an Insight (Newport Spectraphysics) 2-photon laser with two excitation lines of which one is fixed and one tunable (1045nm and 680-1300nm respectively), 6 non-descanned external detectors including 2 nose-piece detectors (GaASP). Images were acquired using a 20x, 1.0N.A., water immersion lens with 1.4mm working distance.

A.1.2 Object detection and clustering

The Imperial College High Performance Computing (HPC) cluster we used contains 8 CPUs and one P1000 GPU with 96 Gb of RAM and was accessed using a conventional laptop with 16GB RAM, Intel i9-9880H CPU with a Quadro T2000 Mobile GPU.

A YOLO-V5X model (`https://github.com/ultralytics/yolov5`) was used as the backbone of the 2D object-detection neural network. More details on this model, and on convolutional neural networks more generally, can be found in the literature [6, 133, 70, 123]. In this section we provide only a brief overview of the model with a focus on the manner in which the algorithm identifies objects; and how it was adapted to generate 3D estimates. Image transformations, such rotation or artificial introduction of additional noise, are only used in the training and validation stages to improve the neural network's performance when applied to unaltered experimental sample images.

The YOLO algorithm works by placing a multitude of boxes within the space of a 2D image and then filtering these boxes based on probability estimates from the model. It is a fully connected neural network which divides a 2D image into $S \times S$ grid of cells into which B bounding boxes are detected. The model identifies a set of box sizes for each class *a priori* using a k-means clustering algorithm run on box sizes observed within the training data. Each bounding box is defined by 5 parameters: the x, y central position, width (w), height (h) and a confidence score, C. This last value, C, is the confidence estimate over the presence, or absence, of an object being within the grid cell. This makes use of the intersection-over-union (IOU) between a predicted bounding box and a ground

truth (manually annotated) bounding box. The greater the overlap between the two, the higher the IOU, and the greater the confidence in the box. If any object is absent from the grid cell, the probability of the object ($Pr(\text{Object})$) is set to 0. Otherwise, it is 1. For the i^{th} bounding box in the j^{th} grid cell, the confidence score, C_{ij} is thus calculated as [123]:

$$C_{ij} = P(\text{Object}_{ij}) \times \text{IOU}$$

In addition to these five parameters, a set of conditional class probabilities is calculated. Given K possible classes, this is the probability of the object belonging to any specific k^{th} class: $\Pr(\text{Class}_k|\text{Object})$. A class-specific confidence score (CS_{kij}) is then calculated as a product of C_{ij} and the conditional class probability [123]:

$$\begin{aligned} CS_{ijk} &= P(\text{Object}_{ij}) \times \text{IOU} \times \Pr(\text{Class}_{kij}|\text{Object}_{ij}) \\ &= P(\text{Class}_{ijk}) \times \text{IOU} \end{aligned}$$

The class-specific confidence scores and IOU results are used to *select* bounding boxes through *non-maximum suppression* (NMS) [15, 123]. YOLOv5 makes use of *soft*-NMS which is better adapted to overlapping objects [15].

To aggregate the final set of 2D bounding boxes into 3D bounding 'cubes' we ordered the bounding boxes for each class by maximum diameter and mean fluorescence intensity (mFI). For each box within the set of boxes (B) within the k^{th} class, starting from the largest and brightest boxes, a central x, y, z location is calculated, which we call q. We can also determine a maximum diameter for this box, d. From this point q the surrounding cluster of boxes in the z dimension which have a distance from q which is $< d/2$. We call this set of clustered bounding boxes B_c. Within this context, $B_c \subset B$ but all the 2D bounding boxes within B_c are assumed to belong to a single cell (cube) surrounding an individual cell. Once identified, B_c is removed from the B. The process is repeated until every 2D box is allocated to a 3D cube.

To apply this model, an $x \times y \times z \times c$ dimensional image, where $c = 3$ for RGB, was divided into a set of $416 \times 416 \times 1 \times 3$ (RGB) tiles. Test/validate/train subsets were selected from a random sampling of this tile set. Manual annotation was performed to identify cells of interest within these selected images with a minimum of 500 cells annotated within each cell class. Once trained, the final object detection was performed on

the full set of tiles.

A.2 IVM experiments

A.2.1 Mice

All animal work was performed in accordance with the UK Home Office regulations under the Animals (Scientific Procedures Act) 1986 and the Animal Welfare and Ethical Review Body (AWERB) guidelines at Imperial College London and Sir Francis Crick Institute. All procedures were approved by the UK Home Office under project license PP9504145.

All mice were bred and housed at Imperial College London or Sir Francis Crick Institute, according to institutional guidelines. mT/mG mice [107] were obtained from Sir Francis Crick Institute. Flk-1-GFP mice [154] were a gift from Alexander Medvinsky (University of Edinburgh); Nestin-GFP mice from Paul Frenette (Einstein School of Medicine, New York) [105]; Flk2-cre mice were a gift from Camilla Forsberg (University of California, Santa Cruz) [12, 16]. For all experiments and transplantation assays, cohorts of male mice between 11 to 16 weeks were used.

A.2.2 HSC and MPP transplantation

All transplantation experiments were carried out under strict sterile conditions. Bones from healthy mT/mG or mT/mG x Flk2-cre donor mice were harvested (femurs, tibias, iliac bones, sternum, and vertebrae), gently crushed in FACS buffer (2% Foetal Bovine Serum (FBS; Life Technologies, Cat Nº 10500-064) in Dulbecco's Phosphate Buffered Saline, without calcium chloride and magnesium chloride (PBS; Sigma Aldrich, Cat NºD8537)) using a mortar and a pestle, pipetted into a single cell suspension, and filtered through a 40μm strainer. The samples were then centrifuged at 500g for 5 minutes at 4ºC, resuspended and incubated for 5 minutes in red blood cell lysis buffer (0.001g/ml Potassium bicarbonate (Cat Nº 60339), 0.008g/ml Ammonium chloride (Cat Nº A9434), 20mM EDTA (Cat Nº EDS) (all from Sigma), and 5% (v/v) FBS in Milli-Q water) at room temperature. After that time, samples were centrifuged again at 500xg for 5 minutes at 4ºC. The resulting single cell suspensions were stained with the biotinylated lineage cocktail (CD3, CD4, CD8, Ter119, B220, Gr-1 and CD11b). Samples were lineage depleted using streptavidin magnetic microbeads (130-048-101; Miltenyi Biotech), LD columns (130-042-901; Miltenyi Biotech) and magnets (Miltenyi Biotech). The lineage depleted samples were then centrifuged and stained with the following antibody panel: Fixable Viability 780, Streptavidin BV510, C-Kit PerCP Cy5.5, Sca1 BV711, CD48 PE Cy7, CD150 BV650, CD34 APC, Flk2 CF-594.

All samples were cell sorted using a FACSAria III (BD Biosciences). The sorted cells were transplanted in recipient mice that were conditioned with lethal irradiation administered in two doses of 5.5Gy. 10,000 to 12,000 sorted cells were transplanted intravenously into each mouse, 3-4 hours after irradiation.

A.2.3 Intravital microscopy

IVM was performed using an LSM 780 Zeiss upright confocal/two photon hybrid microscope equipped with Argon laser (458, 488 and 514 nm), a diode-pumped solid-state 561 nm laser, a Helium-Neon laser (633 nm), and tuneable infrared multiphoton laser (Spectraphysics Mai Tai DeepSee 690-1020 nm), 4 non-descanned detectors (NDD) and an internal spectral detector array. The signal was visualised using a Zeiss W Plan-Apochromat x20 DIC water immersion lens (1.0 N.A.). An external dichroic mirror (450nm) was used when confocal and two photon microscopy were combined. Surgery and BM calvarium intravital imaging were performed as previously described [40, 58, 62, 89, 119]. Mice were anaesthetised using isoflurane in medical oxygen, using 4% isoflurane in 2L/min O2 for induction and 2-1% isoflurane and 2L/min O2 for maintenance. Minimal invasive surgery was performed to remove the skin on top of the skull (calvarium) and expose the skull bone. Using Diamond Carve Dental Cement (Associated Dental Products; Cat. Nº SUN527) a small custom-made metal headpiece (imaging window) was attached to the bone. Once the cement was completely solidified, the mouse was transferred to the microscope stage. The mouse was placed on a heated pad and its temperature was constantly monitored through a thermal probe connected to a thermal sensor. To protect the eyes from dryness and corneal injury a lubricant eye ointment was applied (Lacri-lube®). The lock and key mechanism linking the headpiece to the mouse holder positioned onto the stage, allowed continuous stable imaging of the calvarium BM. A small amount of PBS was placed on the window to soften the subcutaneous tissue on top of the bone, which was then cleaned with a cotton bud. PBS was added for imaging acquisition. Blood vessels were visualised using 80μl of 8mg/ml Cy5-Dextran (500KDa; NANOCS, DX500-S5-1) injected intravenously. Second harmonic signal was used to visualise the bone and was excited with the two-photon laser tuned at 870nm and detected with external detectors. GFP signal was excited at 488nm, tdTomato at 561nm and Cy5 was excited with 639nm, and all of them were detected using the internal detectors.

Acquisition of 3D tile scans of the whole BM area within the imaging window, defined by coronal suture, central sinus, bifurcation, and bone were done by stitching of adjacent z-stack images. The tile scans allowed the visualisation and analysis of calvarium BM at

single cell resolution. Acquiring specific positions of interest within the tile scan repeatedly over time (3 minutes interval between timeframes) allowed to produce time lapse datasets to study cell dynamics and interactions.

When it was necessary to image the same positions during different imaging sessions, mice were subcutaneously injected with analgesia (Buprenorphine 0.1mg/kg) one hour previous to the end of the first imaging session. The imaging window was covered with intrasite gel (Intrasite gel applipak (8g), Smith&Nephew, Cat Nº 66027308) to protect the bone and prevent scar formation, and bandaged. Mice were allowed to recover from anaesthesia and placed in a separate cage, where buprenorphine mixed in raspberry jelly (0.8mg/kg) was provided. Mice could then be re-anaesthetized, placed on the microscope stage and the same positions could be found and re-imaged.

A.2.4 Flow cytometry

Samples were first stained for 20 minutes at 4ºC in the dark with a lineage antibody cocktail (1:40), which includes the following biotinylated antibodies: CD3, CD4, CD8, Ter119, B220, Gr-1 and sometimes CD11b mixed 1:1 ratio. After staining, the cells were washed in FACS buffer and centrifuged at 500g for 5 minutes. The sample was then resuspended in an antibody mix prepared in FACS buffer, containing all conjugated antibodies in Table A.1 and streptavidin (1:1000). All conjugated antibodies were used 1:200 apart from CD34 (1:80) and Flk2 (1:100). After the staining, the cells were washed in FACS buffer, centrifuged at 500xg for 5 minutes at 4ºC and resuspended in FACS buffer for its acquisition. 4,6-diamidino-2-phenylindole (DAPI) (NucBlue® Fixed Cell ReadyProbes® Reagent; Invitrogen, Cat. # R37606) was used to assess viability and separate live and dead cells.

Single colour controls for compensation were prepared using OneComp eBeads (01-1111-42; Biolegend) or WT BM cells. Fluorescence-minus-one (FMOs) controls were prepared when necessary, using sample cells and following the same staining protocol.

For the calculation of absolute cell numbers Calibrite Beads (BD Biosciences, Cat. No. 340486) were used, as previously described [61]. Briefly, calibrate beads were suspended in PBS at a concentration of 1×10^6 beads/ml and mixed using a 10ml pipette to ensure a homogeneous distribution of beads. From this stock 100μl containing 100,000 beads were added to the sample before acquisition so that beads would represent 1-10% of the total number of mononuclear cells contained in the sample. Final absolute cell numbers were back calculated using the number of acquired beads, acquired cells and the concentration

of beads added to each sample.

For the analysis of peripheral blood reconstitution, 50-100 μl of peripheral blood were collected by venipuncture into EDTA coated tubes. Red blood cells were lysed as described above, and cells were stained with the following antibodies: CD45, CD11b, CD3 and B220.

All the samples were acquired with a LSR Fortessa (BD Biosciences), and the data were analysed with FlowJo (Tree Star) software.

A.2.5 Wholecount staining

Calvaria were collected from mice after IVM sessions and fixed overnight in 4% PFA. Samples were always kept at 4ºC, incubated in a blocking buffer for 24 hours, then with primary antibodies for 48 hours, washed in PBS, incubated with secondary antibodies for 24 hours and washed again. They were kept in PBS for microscopy analysis, which were performed using the same settings as for intravital microscopy.

A.2.6 Image processing and quantification

ZEN Black (Zeiss, Germany) software was used to stitch three-dimensional BM tile scans.

Cell tracking was performed manually using IMARIS (Bitplane, Switzerland). To increase accuracy in cell tracking data, displacements in the Z plane caused by movement artifacts (e.g. breathing) was corrected by applying 4D data registration protocols implemented in FIJI (Preibisch et al., 2010) that allowed to register the time-lapses before cell tracking. Coordinates from each cell at each time point were exported from IMARIS for further analysis.

Buffer	Recipe
FACS Buffer	2% Foetal Bovine Serum (FBS; Life Technologies, Cat Nº 10500-064) in Dulbecco's Phosphate Buffered Saline, Modified, without calcium chloride and magnesium chloride (PBS; Sigma Aldrich, Cat NºD8537; or PBS tablets, Gibco, Cat Nº 18912-014)
Red Blood Lysis Buffer	0.001g/ml Potassium bicarbonate (Cat Nº 60339), 0.008g/ml Ammonium chloride (Cat Nº A9434), 20mM EDTA (Cat Nº EDS) (all from Sigma), and 5% (v/v) FBS in Milli-Q water

Table A.1: Flow cytometry.

A.3 DivisionCounter experiments

A.3.1 DivisionCounter construct cloning

To construct the DivisionCounter backbone, the pLVX-EF1α-AcGFP1-N1 plasmid (Clontech 631983) was amplified using primers targeting *EcoR1* restriction site and the start of *puroR* gene, in order to remove *AcGFP1* and *PGK* promotor from the plasmid. *mScarlet* gene without start codon followed by a P2A element upstream of puroR was added using pmScarlet_C1 (Addgene #85042) plasmid as template. A two-site directed mutagenesis to insert a *Nde1* restriction site downstream puroR and a *Spe1* restriction site upstream of *EcoR1* was performed. In the end we obtained the EF1alpa-Spe1-EcoR1-mScarlet-P2A-puroR-Nde1-WPRE plasmid, that was used as our reference DivisionCounter backbone. Small ($\leq$ 100bp) oligonucleotides were synthetized by Eurofins, of the form Spe1-Kozak-Microsatellite-EcoR1, where the Microsatellite was either a guanine repeat (23G, 26G) or a spacer sequence made of 26 nucleotides (26N). All oligonucleotide sequences are supplied in Table A.2. Those small oligonucleotides were hybridized and subsequently ligated into the DivisionCounter backbone previously digested by Spe1(NEB) and EcoR1(NEB) using T4 ligase (NEB). The resulting ligation reaction was added to JM109 competent bacteria that were subsequently grown in SOC medium before being plated onto agar plate containing ampicillin and incubated overnight at 30°C. Clones of interest were then selected, amplified, sanger sequenced to verify the correct ligation of our oligomers and one of those clones was then subsequently amplified in maxi preps to generate our final plasmid stock. All oligos used are in Table A.2.

A.3.2 Lentivirus production, transduction and titration

HEK293T cells were plated at seeding density of 12×10^6 in T150 in DMEM glutamax (Gibco 61965-026) supplemented with 10% FCS (Dutscher S1900-500C), 1% sodium pyruvate (Gibco, 11360070) and 1% MEM-NEAA(Gibco, 11140050) and grown overnight at 37°C with 5% CO_2. The following day, the medium was replaced, and cells were transfected with the DivisionCounter plasmid, the psPAX2 lentiviral packaging plasmid (Addgene plasmid #12260,) and the pMD2.G envelope plasmid (Addgene, plasmid #12259). Using a 2.5:1 ratio of psPAX2:pMD2G, the packaging and envelope plasmids were incubated in a NaCl solution (150 mM) containing 1:12 Polyethylenimine "Max" (PEI MAX; polyScience #24765, 7.5 mM) for 15 minutes. Two days after transfection, supernatant was collected and filtered through a 0.45 μm filter and concentrated using Amicon Ultra-15 filters (100 kDa) (Millipore #UFC9100) via centrifugation. Viruses were then eluted in Opti-MEM (Gibco, 31985062) and stored at -80 °C until further use.

DivisionCounter	Oligo Name	Oligo Sequence
Control (26N)	**FW-26ctl-Spe1**	CTAGTCAGATCTCGAGCTCA AGCTTGCCACCATGGGGATC CGAGGAGATCCACCGGTCGCC ACGGGATCAGGATCAGGATTCG
	Rev-26ctl-Spe1	AATTCGAATCCTGATCCTGA TCCCGTGGCGACCGGTGGAT CTCCTCGGATCCCCATGGTGG CAAGCTTGAGCTCGAGATCTGA
26G	**FW-26G-Spe1**	CTAGTCAGATCTCGAGCTCA AGCTTGCCACCATGGGGATC CGGGGGGGGGGGGGGGGGGGG GGGGGGTCAGGATCAGGATTCG
	Rev-26G-Spe1	AATTCGAATCCTGATCCTGA CCCCCCCCCCCCCCCCCCCC CCCCCCGGATCCCCATGGTGG CAAGCTTGAGCTCGAGATCTGA
23G	**FW-23G-Spe1**	CTAGTCAGATCTCGAGCTCA AGCTTGCCACCATGGGGATC CGGGGGGGGGGGGGGGGGGGG GGGTCAGGATCAGGATTCG
	Rev-23G-Spe1	AATTCGAATCCTGATCCTGA CCCCCCCCCCCCCCCCCCCC CCCGGATCCCCATGGTGGCAA GCTTGAGCTCGAGATCTGA

Table A.2: Oligo sequence used for cloning.

The viral titter and concentration were determined for each cell line by titrating the viral stock on 5×10^5 newly plated cells (at 0.5×10^5 cells/mL in a 24-wells plate) in their specific culture medium and measuring mScarlet expression 3 days post infection by flow cytometry with Cytoflex LX (Beckman). A MOI of 0.1 was targeted to ensure the majority (over 90%) of the infected cells would have been successfully transduced with only one construct.

A.3.3 DivisionCounter experiment in vitro

Culture medium and DivisionCounter infection of HEK293T cells

Human embryonic kidney 293 T (HEK293T) cells were cultured at 37°C, 5% CO2 in Dulbecco's Modified Eagle's Medium (DMEM+GlutaMAX; Gibco 61965-026), supplemented with 10% heat inac5vated Fetal Calf Serum (FCS; Dutscher S1900-500C) and 1% Penicillin Streptomycin (Pen Strep; Gibco 15140-122).

HEK293T were transduced with DivisionCounter viruses at MOI 0.1. 2 to 6 days later, HEK293T cells were collected, stained with SYTOX Green (Invitrogen, R37168, 1/40 dilution, incubated no more than 30 min before acquisition) and live mScarlet-positive cells were sorted on BD ARIA FUSION (BD Bioscience). 25000 sorted mScarlet+ HEK293T were subsequently plated in a 24-wells plate and incubated at 37°C or 33°C with 5% CO_2. Cells were split twice a week for up to 46 days post sort by removing the cells supernatant, incubating with TrypLE Express (Gibco, 12605010) at 37°C, 5% CO_2 before resuspending in full medium, counted and plated back in a fresh plate or flask.

Average Cell Division Count from cell count measurement (cell population doubling)

At each acquisition time point, before each splitting, we counted the absolute number of cells in each well using the Scepter 3.0 handheld automated Cell counter (Merck, PHCC360KIT, PHCC360250). To obtain the cell population division count over time, we computed the cell population doublings (PD) at each time point (t) using the absolute cell counts:

$$PD = log_2(\frac{\frac{n_t}{\prod_t^{i=0} d_i}}{n_0})$$

where n_t is the absolute number of cells measured at each time point, d_i is the inverse of the cumulative product of splitting ratios, and n0 is the number of cells initially plated. Assuming cells grew exponentially with no cell death; the cell population doubling then equals the average division count of the cell population.

m-scarlet measurements for the DivisionCounter read-out on HEK293T

When splitting HEK293T cells, left over HEK293T cells were washed and stained with SYTOX Green Green (Invitrogen, R37168, 1/40 dilution, incubated no more than 30 min before acquisition). mScarlet expression was measured with Cytoflex LX (Beckman) and the percentage of mScarlet-positive cells over time were analysed with FlowJo (BD Bioscience). This percentage was used either for the statistical inference of the fluorescence loss probability or to compute the average division count with the DivisionCounter formula.

mScarlet reacquisition experiment with HEK293T

In order to track mScarlet reacquisition in HEK293T, we followed a similar protocol as detailed before except that we sorted mScarlet-negative cells from sorted mScarlet-positive cells that were left in culture for 20 to 40 days post sort and that had lost their mScarlet expression in the meantime/as they divided. Cytometry analysis, cell counting and cell population doubling computation were performed as before.

In vitro experiments on other cell types (MEF and Nalm6)

We performed similar *in vitro* experiments as with HEK293T cells on Mouse Embryonic Fibroblast (MEF), and Nalm6 cells. MEF were cultured in the same medium as HEK293T cells, with inactivated FCS (Gibco 15140-122). Human Nalm6 cells, a B cell precursor leukaemia cell line, expressing firefly luciferase (Fluc) and blue fluorescent protein (BFP), called Nalm6 in this paper, were cultured in Roswell Park Memorial Ins5tute (RPMI) 1640 Medium with GlutaMAX^{TM} Supplement (Gibco, 61870-010), supplemented with 10% heat inac5vated Fetal Calf Serum (FCS; Dutscher S1900-500C, Batch: S00CH) and 1% Penicillin Streptomycin (Pen Strep; Gibco 15140-122) at 37°C, 5% CO_2.

After transduction and sorting of mScarlet+ cells and MEF were plated at 5×10^4 cells/ml seeding density and Nalm6 cells at 106 cells/ml seeding density and cultured all at 37°C, 5% CO_2. Every 2-4 day, cells were split and, at each acquisition time point, cells were replated at 105 cells/mL for MEF and between $1 - 3 \times 10^6$ cells/mL for Nalm6. Nalm6 cells being more sensitive to sort, we sorted them at room temperature and collected them in pure Fetal Calf Serum (FCS; Dutscher S1900-500C, Batch: S00CH) before replacing them in culture medium.

When splitting MEF and Nalm6 cells, left over cells were washed and stained with SYTOX Green Green (Invitrogen, R37168, 1/40 dilution, incubated no more than 30 min before acquisition). mScarlet expression was measured with Cytoflex LX (Beckman) and the per-

centage of m-Scarlet-positive cells over time were analysed with FlowJo (BD Bioscience) as for HEK293T cells.

A.3.4 DivisionCounter experiment in vivo

Mice

All mouse studies were carried out in accordance with guidelines and approval of French authorities. All in vivo experiments were conducted using 8- to 16-week-old NOD/SCID/IL-2Rγ- null (NSG) female mice purchased from Charles River Laboratories (France), housed in a facility with 12-h light/12-h dark cycle at 22 $\pm$ 1 °C and 40 $\pm$ 10% humidity. Teklad Global 18% protein rodent diet and tap water were provided ad libitum. All mouse experiments were performed using protocols approved by the French Ministry of Higher Education, Research and Innovation. Ethics committee number 118 (APAFIS#26128-2019032323036339, DAP2018-014 and APAFIS#28212-2020111714259701, DAP2020-021).

CAR-T cell production

Buffy coats from anonymous healthy donors were obtained from Etablissement Français du Sang (Paris, France) in accordance with INSERM ethical guidelines. According to French Public Health Law (art L 1121-1-1 and art L 1121-1-2), written consent and Institutional Review Board (IRB) approval are not required for human noninterventional studies. Peripheral blood mononuclear cells (PBMCs) were isolated from healthy donor leukapheresis rings by density gradient centrifugation (Lymphoprep, Stemcell 07851) and frozen in CryoStor cell cryopreservation media (Sigma-Aldrich C2874). T lymphocytes were purified using the Pan T cell isolation kit (Miltenyi Biotech 130-096-535) and activated with Dynabeads human TActivator CD3/CD28 (1:1 bead:cell) (ThermoFisher 11131D) in X-VIVO 15 medium (Lonza 02- 053Q) supplemented with 5% human serum (Sigma Aldrich H4522) and 0.5 mM bmercaptoethanol at density of 106 cells/mL. 24 hours after activation, T cells were transduced with lentiviral supernatants of an anti-CD19 (FMC63)-CD8tm-4IBB-CD3ζ CAR construct (rLV.EF1.19BBz, Flash Therapeutics) at MOI 10 in complete X-VIVO with a final concentration of 4 μg/mL of Polybrene. After 16-18 hours the supernatant was almost completely discarded, and cell suspension was completed to a final volume of 2 mL with complete X-VIVO containing 100 IU/ml of rhIL-2 (Peprotech 200-02). For untransduced T cells, the same protocol was followed without lentiviral particles. CAR-T cells were maintained in X-VIVO supplemented with 100 IU/ml of rhIL-2 and 6 days after transduction, cells were subjected to magnetic removal of activation beads and collected to check CAR expression, phenotyping, killing capacity or freezed for later injections into tumor-bearing mice.

CAR-T cells phenotyping was assessed using a CD19 CAR detection reagent (Biotinylated) (Miltenyi Biotech 130-129-550) and incubating for 10 minutes at room temperature. After washing, cells were stained with PE-conjugated streptavidin for 30 minutes at 4°C in PBS with BSA 0.5% and EDTA 2 mM. Samples were acquired on BD LSR-II or BD Fortessa X-20 (BD Biosciences) and analyzed with FlowJo software v.10.6.2 (BD Biosciences).

Cytotoxicity assays

The cytotoxicity of T cells transduced with a CAR was determined by co-culturing in triplicates at the indicated E/T ratio, CAR-T cells (effector cells) with NALM6 Luc and BFP-expressing, 26G-DivisionCounter or 26N-DivisionCounter, (target cells) in a total volume of 100 μl of complete X-VIVO per well using U bottom 96-well plate. Target cells were plated at the same cell density alone to determine the maximal luciferase expression (relative light units (RLU)). Effector cells were plated alone to measure the background. 18 h after culturing the cells at 37°C, 50 μl of luciferase substrate (Perkin Elmer 122799-10) at 0.3 mg/ml in PBS were added to each well and luminescence was detected in a SpectraMax ID3 plate reader (Molecular Devices). Specific Lysis was determined as $(1 - \frac{RLU_{sample} - RLU_{background}}{RLU_{max} - RLU_{background}}) \times 100$ %.

Nalm6 infection and sorting

0.25×10^6 NALM6 BFP-expressing 26G-DivisionCounter cells were injected into 8–16-week-old female NSG mice intravenously by tail vein injection. Three days later 2×10^6 CAR-T cells were thawed and injected via lateral tail vein, defining the day 0 of the experiment. Tumour burden was measured by bioluminescence imaging using Xenogen IVIS Imaging System (PerkinElmer) one or two times per week. Acquired bioluminescence data was analysed using Living Image software (PerkinElmer) and expressed in average radiance (photons/sec/cm2706 /sr).

In vivo cell injection

0.25×10^6 Nalm6 cells were injected into NSG female mice intravenously via lateral tail vein injection. For the CAR-T cells experiment, seven days later 2×10^6 CAR-T cells were unfreezed and injected via lateral tail vein, defining the day 0 of the experiment.

In vivo Luciferase assay

Nalm6 tumour burden was measured by bioluminescence imaging using Xenogen IVIS Imaging System (PerkinElmer) one or two times per week. Acquired bioluminescence

data was analysed using Living Image software (PerkinElmer) and expressed in average radiance (photons/sec/cm2706 /sr).

Mouse tissue processing and cell isolation

At each acquisition time point, blood was first collected from the submandibular vein from each mouse. Then cervical dislocation was performed and both femurs and tibias, lungs, liver, and spleen were collected for each mouse and all organs were processed to obtain cell suspensions before antibody staining.

For each blood sample (200uL), red blood cells (RBC) were lysed by hypotonic shock with distilled water and filtered using $70\mu m$ cell strainer before flow cytometry staining.

Bone marrow cells were isolated from femur and tibia bones of mice by bone flushing for Section 4.4.2 or centrifugation (300gx 30sec-1min) for Section 4.4.3. All BM cells were then washed and stained for flow cytometry analysis, except BM cells from the day 7 of Section 4.4.2 experiment that were first enriched. For the enrichment, cells were filtrated through a $70\mu m$ cell stainer and washed in cold MACS buffer (PBS-0,5% BSA). Up to 140.10^6 cells were resuspended in $80\mu l$ of MACS buffer and incubated 15min at 4°C with $20\mu l$ of Mouse cell depletion Cocktail (Miltenyi 130-104-694). Prior magnetic separation cells concentration was adjusted to 30.106 cells/ml using MACS buffer. Enrichment was done with LS column (Miltenyi #130-042-401) following the manufacturer recommendations. Cells were then washed twice with PBS supplemented with 10%FBS then stained for flow cytometry analysis as the other BM cells. Spleen cells were isolated by crushing spleen through $100\mu m$ nylon mesh cell strainer (fisherbrand 22363549) with the plunger end of a syringe, then washed and stained for flow cytometry analysis according to the antibody panel presented below.

Spleen cells were isolated by crushing spleen through $100\mu m$ nylon mesh cell strainer (fisherbrand 22363549) with the plunger end of a syringe, then washed and stained for flow cytometry analysis according to the an5body panel presented below.

Lungs were cut in small pieces and digested with Liberase TL ((Roche, ref: 05401020001, 0.15 mg/mL) and DNAse (Roche, ref: 11284932001, 1 mg/mL) at 37°C for 30 minutes. With the plunger end of a syringe, digested lungs were mashed through a 100 μm cell strainer and washed to obtain single cell suspensions. Before staining, mononuclear cells were enriched using a density gradient medium (Lymphoprep, STEMCELL technologies, Catalog#07851) and centrifugation at 400xg for 30min at room temperature.

Livers were injected and incubated for 15 minutes at 37°C with dissocia5on buffer, dPBS supplemented with 0,01 mg/ml Collagenase D (Roche 11088858001) and 200U/ml DNAse I (Roche 10104159001), then dissociated using GentleMACS (Miltenyi 130-096-427) 37C_m_LIDK_1 programme. Before staining, mononuclear cells were enriched using a density gradient medium (Lymphoprep, STEMCELL technologies, Catalog#07851) and centrifugation at 400xg for 30min at room temperature.

After cell isolation, cells were either directly stained and acquired Section 4.4.2 or kept in CO2-independant media (Gibco 18045-054) overnight before staining and acquisition Section 4.4.3. All cells were stained as explained below.

Antibody staining and flow cytometry acquisition and analysis

After cell counting using the LUNA-FX^{TM} (Logos biosystems), the total number of cells per sample was distributed into 1 to 10 wells with a maximum of 2×10^6 cells/well. In at least 3 wells/sample $50\mu l$ CountBright^{TM} absolute counting beads (ThermoFisher C36950) were added. Cells were then stained with mCD45, mTer119, hCD19, hHLA-DR for Section 4.4.2, and also hCD3-BV510 (for Section 4.4.3) antibodies (details in Table A.3) in PBS with 10% FBS at 4°C for 20 min in the dark. The live/dead cell marker (SYTOX^{TM} Green Ready Flow^{TM} Reagent Invitrogen^{TM} R37168) at a 1/40 dilution was added no more than 30min at RT before acquisition. For BM enriched samples and Ki67 staining samples (Fig 3), cells were washed twice in PBS then stained with Live/Dead (LIVE/DEAD^{TM} Fixable Near-IR Dead Cell Stain Kit Invitrogen^{TM} L3224) at 1:100 dilution for a maximum of $\tilde{1}06$ cells/$100\mu l$ for 15min at room temperature in the dark. Cells were then washed in PBS with 10%FBS and incubated with hCD19 and hHLA A,B,C antibodies and details in Table A.3 in PBS with 10% FBS at 4°C for 20 min in the dark. The cells were then washed in PBS with 10%FBS and either brought for acquisition for the BM enriched samples, or fixed and stained for Ki67 staining samples.

Antibody	Reference	Clone	Dilution
mCD45 APC-Cy7	Biolegend 103116	30-F11	1/100
mTer119-Fitc	BD 557915	TER-119	1/100
hCD19 BV750	Biolegend 302262	HIB19	1/100
hHLA A,B,C AF700	Biolegend 311438	W9/32	1/50
hCD-BV510	BD 564713	HIT3a	1/200
Ki67-BV510	BD 563462	B56	1/100

Table A.3: Antibodies used for the *in vivo* experiment.

Ki67 staining samples were fixed and permeabilized using the FOXP3 Fix/Perm Buffer Set (Biolegend 421403) following the manufacturer recommendations. Ki-67 an5body staining was then performed in Permeabilization Buffer (from the set) at 4°C for 30 minutes, before cells were washed and resuspended in PBS with 10%FBS. Flow cytometry data were acquired on Cytoflex LX (Beckman) and analysed in FlowJo (BD Bioscience). Nalm6 cells were gated as hHLA-DR$^+$ cells in Section 4.4.2 and Section 4.4.3 except for day 20 data of Section 4.4.2, for which a hHLA-DR$^+$hCD19$^+$ was used. mScarlet-expression was gated as in Section 4.4.2 and Section 4.4.3 and the percentage of mScarlet-positive cells percentage was used for analysis using the DivisionCounter mathematical framework.

Absolute Cell count per sample using CountBright^{TM} absolute counting beads

To retrieve absolute Nalm6 and T cell counts per sample, we used CountBright^{TM} absolute counting beads as per manufacturer recommendation. A fixed volume ($50\mu l$) of CountBright^{TM} beads was added when cells had just been plated in a 96 wells plate after cell isolation. During flow cytometry sample acquisition, the beads were counted along with cells in Section 4.4.2 and Section 4.4.3. Because the total number of beads is known (given by the lot number), the number of cells per acquired sample (the absolute count) is obtained by relating the number of acquired cells to the number of acquired counting beads, then multiplying it by the number of total beads originally plated with the cells. At least 1000 bead events were acquired to assure a statistically significant sampling of sample volume. We then multiplied this absolute cell count per acquired sample, by the ratio this plated sample volume represented of the total cell suspension of the sample, e.g. if we plated one tenth of the sample we acquired with beads, we would multiply the absolute cell count/ acquired sample by ten to obtain the absolute cell count/total sample.

Bibliography

[1] M. Acar, K. S. Kocherlakota, M. M. Murphy, J. G. Peyer, H. Oguro, C. N. Inra, C. Jaiyeola, Z. Zhao, K. Luby-Phelps, and S. J. Morrison. Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature*, 526(7571): 126–130, 2015.

[2] G. Adams, F. S. Tissot, C. Liu, C. Brunsdon, K. R. Duffy, and C. Lo Celso. PACESS: Practical AI-based Cell Extraction and Spatial Statistics for large 3D bone marrow tissue images. *bioRxiv*, 2024. URL `https://www.biorxiv.org/content/early/2024/09/12/2022.12.29.521787`.

[3] J. Adolfsson, R. Månsson, N. Buza-Vidas, A. Hultquist, K. Liuba, C. T. Jensen, D. Bryder, L. Yang, O.-J. Borge, L. A. Thoren, K. Anderson, E. Sitnicka, Y. Sasaki, M. Sigvardsson, and S. E. W. Jacobsen. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential: a revised road map for adult blood lineage commitment. *Cell*, 121(2):295–306, 2005.

[4] O. Akinduro, T. S. Weber, H. Ang, M. Haltalli, N. Ruivo, D. Duarte, N. Rashidi, E. D. Hawkins, K. R. Duffy, and C. Lo Celso. Proliferation dynamics of acute myeloid leukaemia and haematopoietic progenitors competing for bone marrow space. *Nature Communications*, 9(1):1–12, 2018.

[5] M. Allaoui, M. L. Kherfi, and A. Cheriet. Considerably improving clustering algorithms using UMAP dimensionality reduction technique: a comparative study. In *International Conference on Image and Signal Processing*, pages 317–325. Springer, 2020.

[6] N. Aloysius and M. Geetha. A review on deep convolutional neural networks. In *2017 International Conference on Communication and Signal Processing*, pages 0588–0592. IEEE, 2017.

[7] N. Altman and M. Krzywinski. Points of significance: clustering. *Nature Methods*, 14(6):545–547, 2017.

[8] D. A. Arber, A. Orazi, R. Hasserjian, J. Thiele, M. J. Borowitz, M. M. Le Beau, C. D. Bloomfield, M. Cazzola, and J. W. Vardiman. The 2016 revision to the world health organization classification of myeloid neoplasms and acute leukemia. *Blood, The Journal of the American Society of Hematology*, 127(20):2391–2405, 2016.

[9] N. Asada, Y. Kunisaki, H. Pierce, Z. Wang, N. F. Fernandez, A. Birbrair, A. Ma'ayan, and P. S. Frenette. Differential cytokine contributions of perivascular haematopoietic stem cell niches. *Nature Cell Biology*, 19(3):214–223, 2017.

[10] D. M. Barrett, A. E. Seif, C. Carpenito, D. T. Teachey, J. D. Fish, C. H. June, S. A. Grupp, and G. S. Reid. Noninvasive bioluminescent imaging of primary patient acute lymphoblastic leukemia: a strategy for preclinical modeling. *Blood, The Journal of the American Society of Hematology*, 118(15):e112–e117, 2011.

[11] E. Becht, L. McInnes, J. Healy, C.-A. Dutertre, I. W. Kwok, L. G. Ng, F. Ginhoux, and E. W. Newell. Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*, 37(1):38–44, 2019.

[12] C. Benz, V. C. Martins, F. Radtke, and C. C. Bleul. The stream of precursors that colonizes the thymus proceeds selectively through the early T lineage precursor stage of T cell development. *The Journal of Experimental Medicine*, 205(5):1187–1199, 2008.

[13] C. Benz, M. R. Copley, D. G. Kent, S. Wohrer, A. Cortes, N. Aghaeepour, E. Ma, H. Mader, K. Rowe, C. Day, D. Treloar, R. R. Brinkman, and C. J. Eaves. Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. *Cell Stem Cell*, 10(3):273–283, 2012.

[14] J. M. Bernitz, H. S. Kim, B. MacArthur, H. Sieburg, and K. Moore. Hematopoietic stem cells count and remember self-renewal divisions. *Cell*, 167(5):1296–1309, 2016.

[15] N. Bodla, B. Singh, R. Chellappa, and L. S. Davis. Soft-NMS–improving object detection with one line of code. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 5561–5569, 2017.

[16] S. W. Boyer, A. V. Schroeder, S. Smith-Berdan, and E. C. Forsberg. All hematopoietic cells develop from hematopoietic stem cells through Flk2/Flt3-positive progenitor cells. *Cell Stem Cell*, 9(1):64–73, 2011.

[17] K. Bresser, L. Kok, A. C. Swain, L. A. King, L. Jacobs, T. S. Weber, L. Perié, K. R. Duffy, R. J. de Boer, F. A. Scheeren, and T. N. Schumacher. Replicative history marks transcriptional and functional disparity in the CD8+ T cell memory pool. *Nature Immunology*, 23(5):791–801, 2022.

[18] I. Bruns, D. Lucas, S. Pinho, J. Ahmed, M. P. Lambert, Y. Kunisaki, C. Scheiermann, L. Schiff, M. Poncz, A. Bergman, and P. S. Frenette. Megakaryocytes regulate hematopoietic stem cell quiescence through CXCL4 secretion. *Nature Medicine*, 20 (11):1315–1320, 2014.

[19] N. Cabezas-Wallscheid, D. Klimmeck, J. Hansson, D. B. Lipka, A. Reyes, Q. Wang, D. Weichenhan, A. Lier, L. Von Paleske, S. Renders, P. Wünsche, P. Zeisberger,

D. Brocks, L. Gu, C. Herrmann, S. Hass, M. A. G. Essers, B. Brors, R. Eils, W. Huber, M. D. Milsom, C. Plass, J. Krijgsveld, and A. Trumpp. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell*, 15(4):507–522, 2014.

[20] B. Calabrese. Data cleaning. *Encyclopedia of bioinformatics and computational biology: ABC of bioinformatics*, 472, 2018.

[21] J. Cao, H. Cheng, M. Shi, G. Wang, W. Chen, K. Qi, H. Li, J. Qiao, J. Zhao, Q. Wu, L. Zeng, G. Jing, J. Zheng, and K. Xu. Humanized CD19-specific chimeric antigen-receptor T-cells in 2 adults with newly diagnosed B-cell acute lymphoblastic leukemia. *Leukemia*, 33(11):2751–2753, 2019.

[22] K. M. Cappell and J. N. Kochenderfer. Long-term outcomes following CAR T cell therapy: what we know so far. *Nature Reviews Clinical Oncology*, 20(6):359–371, 2023.

[23] J. Carrelha, Y. Meng, L. M. Kettyle, T. C. Luis, R. Norfo, V. Alcolea, H. Boukarabila, F. Grasso, A. Gambardella, A. Grover, K. Högstrand, A. M. Lord, A. Sanjuan-Pla, P. S. Woll, C. Nerlov, and S. E. W. Jacobsen. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*, 554(7690):106–111, 2018.

[24] M. Cazaux, C. L. Grandjean, F. Lemaître, Z. Garcia, R. J. Beck, I. Milo, J. Postat, J. B. Beltman, E. J. Cheadle, and P. Bousso. Single-cell imaging of CAR T cell activity in vivo reveals extensive functional and anatomical heterogeneity. *Journal of Experimental Medicine*, 216(5):1038–1049, 2019.

[25] H. Cheng, Z. Zheng, and T. Cheng. New paradigms on hematopoietic stem cell differentiation. *Protein & Cell*, 11(1):34–44, 2020.

[26] Y. W. Choo, J. Jeong, and K. Jung. Recent advances in intravital microscopy for investigation of dynamic cellular behavior in vivo. *BMB Reports*, 53(7):357, 2020.

[27] C. Christodoulou, J. A. Spencer, S.-C. A. Yeh, R. Turcotte, K. D. Kokkaliaris, R. Panero, A. Ramos, G. Guo, N. Seyedhassantehrani, T. V. Esipova, S. A. Vinogradov, S. Rudzinskas, Y. Zhang, A. S. Perkins, S. H. Orkin, R. A. Calogero, T. Schroeder, C. P. Lin, and F. D. Camargo. Live-animal imaging of native haematopoietic stem and progenitor cells. *Nature*, 578(7794):278–283, 2020.

[28] J. Cosgrove, L. S. Hustin, R. J. de Boer, and L. Perié. Hematopoiesis in numbers. *Trends in Immunology*, 42(12):1100–1112, 2021.

[29] D. L. Coutu, K. D. Kokkaliaris, L. Kunz, and T. Schroeder. Multicolor quantitative confocal imaging cytometry. *Nature Methods*, 15(1):39–46, 2018.

[30] M. J. Cox, F. Lucien, R. Sakemura, J. C. Boysen, Y. Kim, P. Horvei, C. M. Roman, M. J. Hansen, E. E. Tapper, E. L. Siegler, C. Forsman, S. B. Crotts, K. J. Schick, M. Hefazi, M. W. Ruff, I. Can, M. Adada, E. Bezerra, L. A. Kankeu Fonkoua, W. K. Nevala, E. Braggio, W. Ding, S. A. Parikh, N. E. Kay, and S. S. Kenderian. Leukemic extracellular vesicles induce chimeric antigen receptor T cell dysfunction in chronic lymphocytic leukemia. *Molecular Therapy*, 29(4):1529–1540, 2021.

[31] S. Crippa and M. E. Bernardo. Mesenchymal stromal cells: role in the BM niche and in the support of hematopoietic stem cell transplantation. *Hemasphere*, 2(6): e151, 2018.

[32] M. M. Cuadrado, R. M. Szydlo, M. Watts, N. Patel, H. Renshaw, J. Dorman, M. Lowdell, S. Ings, C. Anthias, A. Madrigal, S. Mackinnon, P. Kottaridis, B. Carpenter, R. Hough, E. Morris, K. Thomson, K. S. Peggs, and R. Chakraverty. Predictors of recovery following allogeneic CD34+-selected cell infusion without conditioning to correct poor graft function. *Haematologica*, 105(11):2639, 2020.

[33] H. Dai, Y. Wang, X. Lu, and W. Han. Chimeric antigen receptors modified T-cells for cancer therapy. *Journal of the National Cancer Institute*, 108(7):djv439, 2016.

[34] R. J. De Boer and A. S. Perelson. Quantifying T lymphocyte turnover. *Journal of Theoretical Biology*, 327:45–87, 2013.

[35] M. M. de Jong, L. Chen, M. H. Raaijmakers, and T. Cupedo. Bone marrow inflammation in haematological malignancies. *Nature Reviews Immunology*, pages 1–16, 2024.

[36] N. Dehaen, M. Van Hul, L. Mignion, A.-N. Kouakou, P. D. Cani, and B. F. Jordan. Luciferase transduction and selection protocol for reliable in vivo bioluminescent measurements in cancer research. *Heliyon*, 10(13):e33356, 2024. ISSN 2405-8440.

[37] H. Döhner, E. Estey, D. Grimwade, S. Amadori, F. R. Appelbaum, T. Büchner, H. Dombret, B. L. Ebert, P. Fenaux, R. A. Larson, R. L. Levine, F. Lo-Coco, T. Naoe, D. Niederwieser, G. J. Ossenkoppele, M. Sanz, J. Sierra, M. S. Tallman, H.-F. Tien, A. H. Wei, B. Löwenberg, and C. D. Bloomfield. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood, The Journal of the American Society of Hematology*, 129(4):424–447, 2017.

[38] F. Dong, S. Hao, S. Zhang, C. Zhu, H. Cheng, Z. Yang, F. K. Hamey, X. Wang, A. Gao, F. Wang, Y. Gao, J. Dong, C. Wang, J. Wang, Y. Lan, B. Liu, H. Ema,

F. Tang, B. Göttgens, P. Zhu, and T. Cheng. Differentiation of transplanted haematopoietic stem cells tracked by single-cell transcriptomic analysis. *Nature Cell Biology*, 22(6):630–639, 2020.

[39] S. Doulatov, F. Notta, K. Eppert, L. T. Nguyen, P. S. Ohashi, and J. E. Dick. Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. *Nature Immunology*, 11(7):585–593, 2010.

[40] D. Duarte, E. D. Hawkins, O. Akinduro, H. Ang, K. De Filippo, I. Y. Kong, M. Haltalli, N. Ruivo, L. Straszkowski, S. J. Vervoort, C. McLean, T. S. Weber, R. Khorshed, C. Pirillo, A. Wei, S. K. Ramasamy, A. P. Kusumbe, K. R. Duffy, R. H. Adams, L. E. Purton, L. M. Carlin, and C. Lo Celso. Inhibition of endosteal vascular niche remodeling rescues hematopoietic stem cell loss in AML. *Cell Stem Cell*, 22(1):64–77, 2018.

[41] D. Duarte, E. D. Hawkins, and C. Lo Celso. The interplay of leukemia cells and the bone marrow microenvironment. *Blood, The Journal of the American Society of Hematology*, 131(14):1507–1511, 2018.

[42] K. R. Duffy, C. J. Wellard, J. F. Markham, J. H. Zhou, R. Holmberg, E. D. Hawkins, J. Hasbold, M. R. Dowling, and P. D. Hodgkin. Activation-induced B cell fates are selected by intracellular stochastic competition. *Science*, 335(6066):338–341, 2012.

[43] K. R. Duffy, G. Meli, and S. Shneer. The variance of the average depth of a pure birth process converges to 7. *Statistics & Probability Letters*, 150:88–93, 2019.

[44] G. Espindola, G. Câmara, I. Reis, L. Bins, and A. Monteiro. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. *International Journal of Remote Sensing*, 27(14):3035–3040, 2006.

[45] E. Estey and H. Döhner. Acute myeloid leukaemia. *The Lancet*, 368(9550):1894–1907, 2006.

[46] K. Foster, F. Lassailly, F. Anjos-Afonso, E. Currie, K. Rouault-Pierre, and D. Bonnet. Different motile behaviors of human hematopoietic stem versus progenitor cells at the osteoblastic niche. *Stem Cell Reports*, 5(5):690–701, 2015.

[47] A. S. Fotheringham, C. Brunsdon, and M. Charlton. *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons, 2003.

[48] K. J. Friston. Functional and effective connectivity: a review. *Brain Connectivity*, 1(1):13–36, 2011.

[49] J. Fujisaki, J. Wu, A. L. Carlson, L. Silberstein, P. Putheti, R. Larocca, W. Gao, T. I. Saito, C. L. Celso, H. Tsuyuzaki, D. Côté, M. Sykes, T. B. Strom, D. T. Scadden, and C. P. Lin. In vivo imaging of T_{reg} cells providing immune privilege to the haematopoietic stem-cell niche. *Nature*, 474(7350):216–219, 2011.

[50] N. Goardon, E. Marchi, A. Atzberger, L. Quek, A. Schuh, S. Soneji, P. Woll, A. Mead, K. A. Alford, R. Rout, ChaudhurySalma, A. Gilkes, S. Knapper, K. Beldjord, S. Begum, S. Rose, N. Geddes, M. Griffiths, G. Standen, A. Sternberg, J. Cavenagh, H. Hunter, D. Bowen, S. Killick, L. Robinson, A. Price, E. Macintyre, P. Virgo, A. Burnett, C. Craddock, T. Enver, S. E. W. Jacobsen, C. Porcher, and P. Vyas. Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. *Cancer Cell*, 19(1):138–152, 2011.

[51] I. Gollini, B. Lu, M. Charlton, C. Brunsdon, and P. Harris. GWmodel: an R package for exploring spatial heterogeneity using geographically weighted models. *J. Stat. Softw.*, 63(17):1–50, 2015.

[52] A. Gomariz, P. M. Helbling, S. Isringhausen, U. Suessbier, A. Becker, A. Boss, T. Nagasawa, G. Paul, O. Goksel, G. Székely, S. Stoma, S. F. Nørrelykke, M. G. Manz, and C. Nombela-Arrieta. Quantitative spatial analysis of haematopoiesis-regulating stromal cells in the bone marrow microenvironment by 3D microscopy. *Nature Communications*, 9(1):2532, 2018.

[53] A. Gomariz, T. Portenier, C. Nombela-Arrieta, and O. Goksel. Probabilistic spatial analysis in quantitative microscopy with uncertainty-aware cell detection using deep bayesian regression. *Science Advances*, 8(5):eabi8295, 2022.

[54] T. Grinenko, A. Eugster, L. Thielecke, B. Ramasz, A. Krüger, S. Dietz, I. Glauche, A. Gerbaulet, M. Von Bonin, O. Basak, H. Clevers, T. Chavakis, and B. Wielockx. Hematopoietic stem cells can differentiate into restricted myeloid progenitors before cell division in mice. *Nature Communications*, 9(1):1898, 2018.

[55] T. Gueta and Y. Carmel. Quantifying the value of user-level data cleaning for big data: A case study using mammal distribution models. *Ecological Informatics*, 34: 139–145, 2016.

[56] G. Guzman, M. R. Reed, K. Bielamowicz, B. Koss, and A. Rodriguez. CAR-T therapies in solid tumors: opportunities and challenges. *Current Oncology Reports*, 25(5):479–489, 2023.

[57] S. Haas, J. Hansson, D. Klimmeck, D. Loeffler, L. Velten, H. Uckelmann, S. Wurzer, Á. M. Prendergast, A. Schnell, K. Hexel, R. Santarella-Mellwig, S. Blaszkiewicz,

A. Kuck, H. Geiger, M. D. Milsom, L. M. Steinmetz, T. Schroeder, A. Trumpp, J. Krijgsveld, and M. A. Essers. Inflammation-induced emergency megakaryopoiesis driven by hematopoietic stem cell-like megakaryocyte progenitors. *Cell Stem Cell*, 17(4):422–434, 2015.

[58] M. L. Haltalli, S. Watcham, N. K. Wilson, K. Eilers, A. Lipien, H. Ang, F. Birch, S. G. Anton, C. Pirillo, N. Ruivo, M. L. Vainieri, C. Pospori, R. E. Sinden, T. C. Luis, J. Langhorne, K. R. Duffy, B. Göttgens, A. M. Blagborough, and C. Lo Celso. Manipulating niche composition limits damage to haematopoietic stem cells during plasmodium infection. *Nature Cell Biology*, 22(12):1399–1410, 2020.

[59] D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100(1):57–70, 2000.

[60] F. Hao, C. Sholy, C. Wang, M. Cao, and X. Kang. The role of T cell immunotherapy in acute myeloid leukemia. *Cells*, 10(12):3376, 2021.

[61] E. D. Hawkins, M. Hommel, M. L. Turner, F. L. Battye, J. F. Markham, and P. D. Hodgkin. Measuring lymphocyte proliferation, survival and differentiation using CFSE time-series data. *Nature Protocols*, 2(9):2057–2067, 2007.

[62] E. D. Hawkins, D. Duarte, O. Akinduro, R. A. Khorshed, D. Passaro, M. Nowicka, L. Straszkowski, M. K. Scott, S. Rothery, N. Ruivo, K. Foster, M. Waibel, R. W. Johnstone, S. J. Harrison, D. A. Westerman, H. Quach, J. Gribben, M. D. Robinson, L. E. Purton, D. Bonnet, and C. Lo Celso. T-cell acute leukaemia exhibits dynamic interactions with bone marrow microenvironments. *Nature*, 538(7626): 518–522, 2016.

[63] L. Hayflick. The limited in vitro lifetime of human diploid cell strains. *Experimental Cell Research*, 37(3):614–636, 1965.

[64] S. Hugues, L. Fetler, L. Bonifaz, J. Helft, F. Amblard, and S. Amigorena. Distinct T cell dynamics in lymph nodes during the induction of tolerance and immunity. *Nature Immunology*, 5(12):1235–1242, 2004.

[65] L. S. Hustin, C. Conrad, C. Liu, J. Fuentealba, S. Menegatti, S. Shneer, A. Battistella, F. Tabarin, T. S. Weber, S. Amigorena, K. R. Duffy, and L. Perié. The DivisionCounter, a method for counting large ranges of cell divisions in vivo, reveals cell dynamics of leukemic cell killing via CAR-T therapy. *bioRxiv*, pages 2024–06, 2024. URL `https://www.biorxiv.org/content/early/2024/06/28/2024.06.24.600543`.

[66] A. Ivanovs, S. Rybtsov, E. S. Ng, E. G. Stanley, A. G. Elefanty, and A. Medvinsky. Human haematopoietic stem cell development: from the embryo to the dish. *Development*, 144(13):2323–2337, 2017.

[67] O. Kamenyeva, C. Boularan, J. Kabat, G. Y. Cheung, C. Cicala, A. J. Yeh, J. L. Chan, S. Periasamy, M. Otto, and J. H. Kehrl. Neutrophil recruitment to lymph nodes limits local humoral response to Staphylococcus aureus. *PLoS pathogens*, 11 (4):e1004827, 2015.

[68] A. S. Kanate, N. S. Majhail, B. N. Savani, C. Bredeson, R. E. Champlin, S. Crawford, S. A. Giralt, C. F. LeMaistre, D. I. Marks, J. L. Omel, P. J. Orchard, J. Palmer, W. Saber, P. A. Veys, P. A. Carpenter, and M. Hamadani. Indications for hematopoietic cell transplantation and immune effector cell therapy: guidelines from the American Society for Transplantation and Cellular Therapy. *Biology of Blood and Marrow Transplantation*, 26(7):1247–1256, 2020.

[69] A. B. Khan, B. Carpenter, P. S. e Sousa, C. Pospori, R. Khorshed, J. Griffin, P. Velica, M. Zech, S. Ghorashian, C. Forrest, S. Thomas, S. Gonzalez Anton, M. Ahmadi, A. Holler, B. Flutter, Z. Ramirez-Ortiz, T. K. Means, C. L. Bennett, H. Stauss, E. Morris, C. Lo Celso, and R. Chakraverty. Redirection to the bone marrow improves T cell persistence and antitumor functions. *The Journal of Clinical Investigation*, 128(5):2010–2024, 2018.

[70] V. Kharchenko and I. Chyrka. Detection of airplanes on the ground using YOLO neural network. In *2018 IEEE 17th international conference on mathematical methods in electromagnetic theory (MMET)*, pages 294–297. IEEE, 2018.

[71] A. Khwaja, M. Bjorkholm, R. E. Gale, R. L. Levine, C. T. Jordan, G. Ehninger, C. D. Bloomfield, E. Estey, A. Burnett, J. J. Cornelissen, D. A. Scheinberg, D. Bouscary, and D. C. Linch. Acute myeloid leukaemia. *Nature Reviews Disease Primers*, 2(1):1–22, 2016.

[72] D. Kobak and P. Berens. The art of using t-SNE for single-cell transcriptomics. *Nature Communications*, 10(1):5416, 2019.

[73] A. Köhler, V. Schmithorst, M.-D. Filippi, M. A. Ryan, D. Daria, M. Gunzer, and H. Geiger. Altered cellular dynamics and endosteal location of aged early hematopoietic progenitor cells revealed by time-lapse intravital imaging in long bones. *Blood, The Journal of the American Society of Hematology*, 114(2):290–298, 2009.

[74] K. D. Kokkaliaris, L. Kunz, N. Cabezas-Wallscheid, C. Christodoulou, S. Renders, F. Camargo, A. Trumpp, D. T. Scadden, and T. Schroeder. Adult blood stem cell

localization reflects the abundance of reported bone marrow niche cell types and their combinations. *Blood, The Journal of the American Society of Hematology*, 136 (20):2296–2307, 2020.

[75] M. Kondo, I. L. Weissman, and K. Akashi. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell*, 91(5):661–672, 1997.

[76] J. Konieczny and L. Arranz. Updates on old and weary haematopoiesis. *International journal of molecular sciences*, 19(9):2567, 2018.

[77] W. Koole, H. S. Schäfer, R. Agami, G. van Haaften, and M. Tijsterman. A versatile microsatellite instability reporter system in human cells. *Nucleic Acids Research*, 41(16):e158–e158, 2013.

[78] S. Kozar, E. Morrissey, A. M. Nicholson, M. van der Heijden, H. I. Zecchini, R. Kemp, S. Tavare, L. Vermeulen, and D. J. Winton. Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. *Cell Stem Cell*, 13(5):626–633, 2013.

[79] A. V. Krivtsov, D. Twomey, Z. Feng, M. C. Stubbs, Y. Wang, J. Faber, J. E. Levine, J. Wang, W. C. Hahn, D. G. Gilliland, T. R. Golub, and S. A. Armstrong. Transformation from committed progenitor to leukaemia stem cell initiated by MLL–AF9. *Nature*, 442(7104):818–822, 2006.

[80] I. Kucinski, J. Campos, M. Barile, F. Severi, N. Bohin, P. N. Moreira, L. Allen, H. Lawson, M. L. Haltalli, S. J. Kinston, D. O'Carroll, K. R. Kranc, and B. Göttgens. A time-and single-cell-resolved model of murine bone marrow hematopoiesis. *Cell Stem Cell*, 31(2):244–259, 2024.

[81] Y. Lai and F. Sun. The relationship between microsatellite slippage mutation rate and the number of repeat units. *Molecular Biology and Evolution*, 20(12):2123–2131, 2003.

[82] E. Laurenti and B. Göttgens. From haematopoietic stem cells to complex differentiation landscapes. *Nature*, 553(7689):418–426, 2018.

[83] E. L. Lehmann, J. P. Romano, and G. Casella. *Testing statistical hypotheses*, volume 3. Springer, 1986.

[84] H. Li, Y. Wang, R. Liu, X. Li, P. Zhang, P. Chen, N. Zhao, B. Li, J. Wang, and Y. Tang. Unraveling resistance mechanisms in anti-CD19 chimeric antigen receptor-T therapy for B-ALL: a novel in vitro model and insights into target antigen dynamics. *Journal of Translational Medicine*, 22(1):482, 2024.

[85] W. Li, R. N. Germain, and M. Y. Gerner. Multiplex, quantitative cellular analysis in large tissue volumes with clearing-enhanced 3D microscopy (Ce3D). *Proceedings of the National Academy of Sciences*, 114(35):E7321–E7330, 2017.

[86] C. Lin, Z. Huang, W. Wen, A. Wu, C. Wang, and L. Niu. Enhancing protein expression in HEK-293 cells by lowering culture temperature. *PloS one*, 10(4): e0123562, 2015.

[87] M. J. Lindstrom and D. M. Bates. Nonlinear mixed effects models for repeated measures data. *Biometrics*, pages 673–687, 1990.

[88] C. Lo Celso and D. T. Scadden. The haematopoietic stem cell niche at a glance. *Journal of Cell Science*, 124(21):3529–3535, 2011.

[89] C. Lo Celso, H. E. Fleming, J. W. Wu, C. X. Zhao, S. Miake-Lye, J. Fujisaki, D. Côté, D. W. Rowe, C. P. Lin, and D. T. Scadden. Live-animal tracking of individual haematopoietic stem/progenitor cells in their niche. *Nature*, 457(7225): 92–96, 2009.

[90] C. Lo Celso, C. P. Lin, and D. T. Scadden. In vivo imaging of transplanted hematopoietic stem and progenitor cells in mouse calvarium bone marrow. *Nature Protocols*, 6(1):1–14, 2011.

[91] B. Lu, C. Brunsdon, M. Charlton, and P. Harris. Geographically weighted regression with parameter-specific distance metrics. *International Journal of Geographical Information Science*, 31(5):982–998, 2017.

[92] S. Lymperi, F. Ferraro, and D. T. Scadden. The HSC niche concept has turned 31: has our knowledge matured? *Annals of the New York Academy of Sciences*, 1192 (1):12–18, 2010.

[93] M. Lynch, M. S. Ackerman, J.-F. Gout, H. Long, W. Sung, W. K. Thomas, and P. L. Foster. Genetic drift, selection and the evolution of the mutation rate. *Nature Reviews Genetics*, 17(11):704–714, 2016.

[94] A. B. Lyons. Analysing cell division in vivo and in vitro using flow cytometric measurement of CFSE dye dilution. *Journal of Immunological Methods*, 243(1-2): 147–154, 2000.

[95] A. B. Lyons, S. J. Blake, and K. V. Doherty. Flow cytometric analysis of cell division by dilution of CFSE and related dyes. *Current Protocols in Cytometry*, 64(1):9–11, 2013.

[96] M. Martino, C. Alati, F. A. Canale, G. Musuraca, G. Martinelli, and C. Cerchione. A review of clinical outcomes of CAR T-cell therapies for B-acute lymphoblastic leukemia. *International Journal of Molecular Sciences*, 22(4):2150, 2021.

[97] G. Mathé, D. E. Thomas, and J. W. Ferrebee. The restoration of marrow function after lethal irradiation in man: a review. *Plastic and Reconstructive Surgery*, 24(4): 407–409, 1959.

[98] J. McDevitt, P. Tierney, and A. Brennan. *Cancer in Ireland 1994-2021: Annual statistical report 2023*. National Cancer Registry, Ireland, 2023.

[99] L. McInnes, J. Healy, and J. Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv:1802.03426*, 10, 2020. URL `https://arxiv.org/abs/1802.03426`.

[100] G. Meli, T. S. Weber, and K. R. Duffy. Sample path properties of the average generation of a Bellman–Harris process. *Journal of Mathematical Biology*, 79:673–704, 2019.

[101] S. Méndez-Ferrer, T. V. Michurina, F. Ferraro, A. R. Mazloom, B. D. MacArthur, S. A. Lira, D. T. Scadden, A. Ma'ayan, G. N. Enikolopov, and P. S. Frenette. Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature*, 466(7308):829–834, 2010.

[102] P. Merli, M. Algeri, F. Del Bufalo, and F. Locatelli. Hematopoietic stem cell transplantation in pediatric acute lymphoblastic leukemia. *Current Hematologic Malignancy Reports*, 14:94–105, 2019.

[103] T. F. Mertens, A. T. Liebheit, J. Ehl, R. Köhler, A. Rakhymzhan, A. Woehler, L. Katthän, G. Ebel, W. Liublin, A. Kasapi, A. Triantafyllopoulou, T. J. Schulz, R. A. Niesner, and A. E. Hauser. MarShie: a clearing protocol for 3D analysis of single cells throughout the bone marrow at subcellular resolution. *Nature Communications*, 15(1):1764, 2024.

[104] K. R. Mesa, K. Kawaguchi, K. Cockburn, D. Gonzalez, J. Boucher, T. Xin, A. M. Klein, and V. Greco. Homeostatic epidermal stem cell self-renewal is driven by local differentiation. *Cell Stem Cell*, 23(5):677–686, 2018.

[105] J. L. Mignone, V. Kukekov, A.-S. Chiang, D. Steindler, and G. Enikolopov. Neural stem and progenitor cells in nestin-GFP transgenic mice. *Journal of Comparative Neurology*, 469(3):311–324, 2004.

[106] J. J. Moré. The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis: proceedings of the biennial Conference held at Dundee, June 28–July 1, 1977*, pages 105–116. Springer, 2006.

[107] M. D. Muzumdar, P.-Y. Chen, K. J. Dorans, K. M. Chung, A. Bhutkar, E. Hong, E. M. Noll, M. R. Sprick, A. Trumpp, and T. Jacks. Survival of pancreatic cancer cells lacking KRAS function. *Nature Communications*, 8(1):1090, 2017.

[108] F. Nakahara, D. K. Borger, Q. Wei, S. Pinho, M. Maryanovich, A. H. Zahalka, M. Suzuki, C. D. Cruz, Z. Wang, C. Xu, P. E. Boulais, A. Ma'ayan, J. M. Greally, and P. S. Frenette. Engineering a haematopoietic stem cell niche by revitalizing mesenchymal stromal cells. *Nature Cell Biology*, 21(5):560–567, 2019.

[109] A. P. Ng and W. S. Alexander. Haematopoietic stem cells: past, present and future. *Cell Death Discovery*, 3(1):1–4, 2017.

[110] C. Nombela-Arrieta, G. Pivarnik, B. Winkel, K. J. Canty, B. Harley, J. E. Mahoney, S.-Y. Park, J. Lu, A. Protopopov, and L. E. Silberstein. Quantitative imaging of haematopoietic stem and progenitor cell localization and hypoxic status in the bone marrow microenvironment. *Nature Cell Biology*, 15(5):533–543, 2013.

[111] J. Orleans-Lindsay, L. Barber, H. Prentice, and M. Lowdell. Acute myeloid leukaemia cells secrete a soluble factor that inhibits T and NK cell proliferation but not cytolytic function–implications for the adoptive immunotherapy of leukaemia. *Clinical & Experimental Immunology*, 126(3):403–411, 2001.

[112] J. H. Park, I. Rivière, M. Gonen, X. Wang, B. Sénéchal, K. J. Curran, C. Sauter, Y. Wang, B. Santomasso, E. Mead, M. Roshal, P. Maslak, M. Davila, R. J. Brentjens, and M. Sadelain. Long-term follow-up of CD19 CAR therapy in acute lymphoblastic leukemia. *New England Journal of Medicine*, 378(5):449–459, 2018.

[113] A. Pelcovits and R. Niroula. Acute myeloid leukemia: a review. *Rhode Island Medical Journal*, 103(3):38–40, 2020.

[114] E. M. Pietras, D. Reynaud, Y.-A. Kang, D. Carlin, F. J. Calero-Nieto, A. D. Leavitt, J. M. Stuart, B. Göttgens, and E. Passegué. Functionally distinct subsets of lineage-biased multipotent progenitors control blood production in normal and regenerative conditions. *Cell Stem Cell*, 17(1):35–46, 2015.

[115] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, S. Heisterkamp, B. Van Willigen, and R. Maintainer. Package 'nlme'. *Linear and nonlinear mixed effects models, version*, 3(1):274, 2017.

[116] J. C. Pinheiro, D. M. Bates, and M. J. Lindstrom. *Model building for nonlinear mixed effects models*. University of Wisconsin, Department of Biostatistics Madison, WI, 1995.

[117] S. Pinho and P. S. Frenette. Haematopoietic stem cell activity and interactions with the niche. *Nature Reviews Molecular Cell Biology*, 20(5):303–320, 2019.

[118] S. Pinho, J. Lacombe, M. Hanoun, T. Mizoguchi, I. Bruns, Y. Kunisaki, and P. S. Frenette. PDGFRα and CD51 mark human nestin+ sphere-forming mesenchymal stem cells capable of hematopoietic progenitor cell expansion. *Journal of Experimental Medicine*, 210(7):1351–1367, 2013.

[119] C. Pirillo, F. Birch, F. S. Tissot, S. G. Anton, M. Haltalli, V. Tini, I. Kong, C. Piot, B. Partridge, C. Pospori, K. Keeshan, S. Santamaria, E. Hawkins, B. Falini, A. Marra, D. Duarte, C. F. Lee, E. Roberts, and C. Lo Celso. Metalloproteinase inhibition reduces aml growth, prevents stem cell loss, and improves chemotherapy effectiveness. *Blood Advances*, 6(10):3126–3141, 2022.

[120] F. Progatzky, M. J. Dallman, and C. Lo Celso. From seeing to believing: labelling strategies for in vivo cell-tracking experiments. *Interface Focus*, 3(3):20130001, 2013.

[121] A. R. Pyzer, D. Stroopinsky, H. Rajabi, A. Washington, A. Tagde, M. Coll, J. Fung, M. P. Bryant, L. Cole, K. Palmer, P. Somaiya, R. K. Leaf, M. Nahas, A. Apel, S. Jain, M. McMasters, L. Mendez, J. Levine, J. Joyce, P. P. Pandolfi, D. Kufe, J. Rosenblatt, and D. Avigan. MUC1-mediated induction of myeloid-derived suppressor cells in patients with acute myeloid leukemia. *Blood, The Journal of the American Society of Hematology*, 129(13):1791–1801, 2017.

[122] B. J. Quah and C. R. Parish. New and improved methods for measuring lymphocyte proliferation in vitro and in vivo using CFSE-like fluorescent dyes. *Journal of Immunological Methods*, 379(1-2):1–14, 2012.

[123] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 779–788, 2016.

[124] Y. Reizel, N. Chapal-Ilani, R. Adar, S. Itzkovitz, J. Elbaz, Y. E. Maruvka, E. Segev, L. I. Shlush, N. Dekel, and E. Shapiro. Colon stem cell and crypt dynamics exposed by cell lineage reconstruction. *PLoS genetics*, 7(7):e1002192, 2011.

[125] C. Ritz and J. C. Streibig. *Nonlinear regression with R*. Springer, 2008.

[126] C. Rius and M. J. Sanz. Intravital microscopy in the cremaster muscle microcirculation for endothelial dysfunction studies. *Methods in Mouse Atherosclerosis*, pages 357–366, 2015.

[127] A. Sanjuan-Pla, I. C. Macaulay, C. T. Jensen, P. S. Woll, T. C. Luis, A. Mead, S. Moore, C. Carella, S. Matsuoka, T. B. Jones, O. Chowdhury, L. Stenson, M. Lutteropp, J. C. A. Green, R. Facchini, H. Boukarabila, A. Grover, A. Gambardella, S. Thongjuea, J. Carrelha, P. Tarrant, D. Atkinson, S.-A. Clark, C. Nerlov, and S. E. W. Jacobsen. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*, 502(7470):232–236, 2013.

[128] P. Schober, C. Boer, and L. A. Schwarte. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5):1763–1768, 2018.

[129] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems (TODS)*, 42(3):1–21, 2017.

[130] J. Secklehner, C. Lo Celso, and L. M. Carlin. Intravital microscopy in historic and contemporary immunology. *Immunology and Cell Biology*, 95(6):506–513, 2017.

[131] J. Seita and I. L. Weissman. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2(6): 640–653, 2010.

[132] R. Sender and R. Milo. The distribution of cellular turnover in the human body. *Nature Medicine*, 27(1):45–48, 2021.

[133] M. J. Shafiee, B. Chywl, F. Li, and A. Wong. Fast YOLO: A fast you only look once system for real-time embedded object detection in video. *arXiv:1709.05943*, 2017. URL `https://arxiv.org/abs/1709.05943`.

[134] G. Shakhar, R. L. Lindquist, D. Skokos, D. Dudziak, J. H. Huang, M. C. Nussenzweig, and M. L. Dustin. Stable T cell–dendritic cell interactions precede the development of both tolerance and immunity in vivo. *Nature Immunology*, 6(7):707–714, 2005.

[135] J. A. Snowden, I. Sánchez-Ortega, S. Corbacioglu, G. W. Basak, C. Chabannon, R. de la Camara, H. Dolstra, R. F. Duarte, B. Glass, R. Greco, A. C. Lankester, M. Mohty, B. Neven, R. P. d. Latour, P. Pedrazzoli, Z. Peric, I. Yakoub-Agha, A. Sureda, and N. Kröger. Indications for haematopoietic cell transplantation for haematological diseases, solid tumours and immune disorders: current practice in europe, 2022. *Bone Marrow Transplantation*, 57(8):1217–1239, 2022.

[136] G. Stegmann, R. Jacobucci, J. R. Harring, and K. J. Grimm. Nonlinear mixed-effects modeling programs in R. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(1):160–165, 2018.

[137] D. Stegner, J. M. vanEeuwijk, O. Angay, M. G. Gorelashvili, D. Semeniak, J. Pinnecker, P. Schmithausen, I. Meyer, M. Friedrich, S. Dütting, C. Brede, A. Beilhack, H. Schulze, B. Nieswandt, and K. G. Heinze. Thrombopoiesis is spatially regulated by the bone marrow vasculature. *Nature Communications*, 8(1):127, 2017.

[138] S. Stevanović, S. R. Helman, J. R. Wunderlich, M. M. Langhan, S. L. Doran, M. L. M. Kwong, R. P. Somerville, C. A. Klebanoff, U. S. Kammula, R. M. Sherry, J. C. Yang, S. A. Rosenberg, and C. S. Hinrichs. A phase II study of tumor-infiltrating lymphocyte therapy for human papillomavirus–associated epithelial cancers. *Clinical Cancer Research*, 25(5):1486–1493, 2019.

[139] A. P. Stone, T. F. Nascimento, and M. N. Barrachina. The bone marrow niche from the inside out: how megakaryocytes are shaped by and shape hematopoiesis. *Blood, The Journal of the American Society of Hematology*, 139(4):483–491, 2022.

[140] E. A. Susaki, K. Tainaka, D. Perrin, H. Yukinaga, A. Kuno, and H. R. Ueda. Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nature Protocols*, 10(11):1709–1727, 2015.

[141] A. Takami. Hematopoietic stem cell transplantation for acute myeloid leukemia. *International Journal of Hematology*, 107(5):513–518, 2018.

[142] T. Terwilliger and M. Abdul-Hay. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood Cancer Journal*, 7(6):e577–e577, 2017.

[143] H.-T. Thai, F. Mentré, N. H. Holford, C. Veyrat-Follet, and E. Comets. Evaluation of bootstrap methods for estimating uncertainty of parameters in nonlinear mixed-effects models: a simulation study in population pharmacokinetics. *Journal of Pharmacokinetics and Pharmacodynamics*, 41:15–33, 2014.

[144] E. D. Thomas, H. L. Lochte Jr, W. C. Lu, and J. W. Ferrebee. Intravenous infusion of bone marrow in patients receiving radiation and chemotherapy. *New England Journal of Medicine*, 257(11):491–496, 1957.

[145] F. S. Tissot, S. Gonzalez-Anton, and C. Lo Celso. Intravital microscopy to study the effect of matrix metalloproteinase inhibition on acute myeloid leukemia cell migration in the Bone Marrow. In *Proteases and Cancer: Methods and Protocols*, pages 211–227. Springer, 2023.

[146] W. R. Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46(sup1):234–240, 1970.

[147] S. Volmering, H. Block, M. Boras, C. A. Lowell, and A. Zarbock. The neutrophil Btk signalosome regulates integrin activation during sterile inflammation. *Immunity*, 44 (1):73–87, 2016.

[148] T. S. Weber, L. Perié, and K. R. Duffy. Inferring average generation via division-linked labeling. *Journal of Mathematical Biology*, 73:491–523, 2016.

[149] A. E. Whiteley, T. T. Price, G. Cantelli, and D. A. Sipkins. Leukaemia: a model metastatic disease. *Nature Reviews Cancer*, 21(7):461–475, 2021.

[150] A. C. Wilkinson, K. J. Igarashi, and H. Nakauchi. Haematopoietic stem cell self-renewal in vivo and ex vivo. *Nature Reviews Genetics*, 21(9):541–554, 2020.

[151] C. Wissler. The Spearman correlation formula. *Science*, 22(558):309–311, 1905.

[152] Q. Wu, J. Zhang, S. Kumar, S. Shen, M. Kincaid, C. B. Johnson, Y. S. Zhang, R. Turcotte, C. Alt, K. Ito, S. Homan, B. E. Sherman, T.-Y. Shao, A. Slaughter, B. Weinhaus, B. Song, M. D. Filippi, H. L. Grimes, C. P. Lin, K. Ito, S. S. Way, J. M. Kofron, and D. Lucas. Resilient anatomy and local plasticity of naive and stress haematopoiesis. *Nature*, 627(8005):839–846, 2024.

[153] Y. Xie, T. Yin, W. Wiegraebe, X. C. He, D. Miller, D. Stark, K. Perko, R. Alexander, J. Schwartz, J. C. Grindley, J. Park, J. S. Haug, J. P. Wunderlich, H. Li, S. Zhang, T. Johnson, R. A. Feldman, and L. Li. Detection of functional haematopoietic stem cell niche using real-time imaging. *Nature*, 457(7225):97–101, 2009.

[154] Y. Xu, L. Yuan, J. Mak, L. Pardanaud, M. Caunt, I. Kasman, B. Larrivée, R. Del Toro, S. Suchting, A. Medvinsky, J. Silva, J. Yang, J.-L. Thomas, A. W. Koch, K. Alitalo, A. Eichmann, and A. Bagri. Neuropilin-2 mediates VEGF-C–induced lymphatic sprouting together with VEGFR3. *Journal of Cell Biology*, 188 (1):115–130, 2010.

[155] O. O. Yeku, T. J. Purdon, M. Koneru, D. Spriggs, and R. J. Brentjens. Armored CAR T cells enhance antitumor efficacy and overcome the tumor microenvironment. *Scientific Reports*, 7(1):10541, 2017.

[156] N. Zacharakis, H. Chinnasamy, M. Black, H. Xu, Y.-C. Lu, Z. Zheng, A. Pasetto, M. Langhan, T. Shelton, T. Prickett, J. Gartner, L. Jia, K. Trebska-McGowan, R. P. Somerville, P. F. Robbins, S. A. Rosenberg, S. L. Goff, and S. A. Feldman.

Immune recognition of somatic mutations leading to complete durable regression in metastatic breast cancer. *Nature Medicine*, 24(6):724–730, 2018.

[157] J. Zhang, Q. Wu, C. B. Johnson, G. Pham, J. M. Kinder, A. Olsson, A. Slaughter, M. May, B. Weinhaus, A. D'Alessandro, J. Douglas Engel, J. X. Jiang, J. M. Kofron, L. F. Huang, V. B. S. Prasath, S. S. Way, N. Salomonis, H. L. Grimes, and D. Lucas. In situ mapping identifies distinct vascular niches for myelopoiesis. *Nature*, 590 (7846):457–462, 2021.