Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M.Sc Adrien Jolly

born in: Tourcoing, France

Oral Examination: 14th of September 2020

An Investigation of the role of the microenvironment in haematopoiesis and T cell development

Referees:

Prof. Dr. Thomas Höfer

Prof. Dr. Ana Martin-Villalba

Abbreviations

BLAST Basic Local Alignment Tool

BM bone marrow

BrdU Bromo Deoxyuridine

CD Cluster of Differentiation

CDR3 Complementarity-Determining-Region 3

DN Double Negative thymocytes

DP Double Positive thymocytes

EdU 5-Ethinyl-2'-deoxyuridin

FACS Fluorescence-Activated Cell Sorting

FCS Fetal Calf Serum

Fig. Figure

FUCCI Fluorescent Ubiquitination-based Cell Cycle Indicator

Il-7 Interleukin 7

Il2rg Interleukin 2 receptor gamma chain

ISP Immature Single Positive

i.p. intra peritoneal

HSC Hematopoietic Stem Cell

HSPC Hematopoietic Stem and Progenitor Cell

MPP Multipotent Progenitor

MSC Mesenchymal Stromal Cell

MHC Major Histocompatiblity Complex

NK cell Natural Killer cell

PBS Phosphate Buffer Saline

PFA Paraformaldehyde

pMHC Major Histocompatiblity Complex+peptide

Rag Recombination activated gene

SCF Stem Cell Factor

TEC Thymic Epithelial Cell

TCR T-Cell Receptor

Summary

The microenvironment is essential to the development of blood cells. Hematopoietic progenitors receive stimuli from neighbouring cells in the form of secreted cytokines or direct contacts via adhesion molecules. These signals regulate the proliferation and differentiation of the developing cells. In this work, we study interactions between haematopoetic progenitors and their microenvironment from different angles and in different biological systems. We first develop a method to identify candidate ligands-receptors interactions based on transcriptomic data, and using this method, we recover known interactions and predict new candidate interactions between hematopoietic stem and progenitor cells and their niche in the bone marrow. Our analysis notably emphasizes the unique role of mesenchymal stem cells and endothelial cells and further reveals that a niche cell population expressing a certain ligand often expresses its antagonist. This finding suggests a dynamic regulation of HSPCs by the niche. We then undertake to characterize quantitatively cell proliferation, a feature which can be regulated by the microenvironment in vivo. We develop a method to determine the length of each cell cycle phase as well as the relative size of the quiescent fraction. We show that our method, which combines a short pulse-chase experiment with EdU, a thymidin analogue, and mathematical modeling, can accurately determine the cell cycle length of a cancer cell line (Tet21N). We apply our method to double positive thymocytes and establish that most cells from this population are quiescent but that a small fraction of these cells is highly proliferative with a mean cell cycle length of 9h, half the time of the Tet21N. In a final part, we study T cell development in the thymus. We study how the murine thymus, an organ which normally depends on influx from bone marrow progenitor can become self sufficient when transplanted in a mouse which does not produce competent bone marrow progenitors. We show that in this context where resident thymocytes do not have to compete for Il-7 with incoming

progenitors, a specific population, the double negative 3 (DN3), can self-renew despite receiving the preTCR signal which normally drives their differentiation. We further show that the longer dwelling time of these DN3 progenitors permits their differentiation into $\gamma\ \delta$ T cells which is normally precluded by the preTCR signal. We show how the limited resources available in the thymus and the flow of incoming progenitor shape the cellular response to the preTCR signal and thereby directly influence the fate of the DN3.

Zusammenfassung

Die zelluläre Mikroumgebung ist ein wesentlicher Faktor für die Entwicklung von Blutzellen. Hämatopoetische Vorläufer erhalten Stimuli von benachbarten Zellen in Form von sekretierten Zytokinen oder direkten Kontakten über Adhäsionsmoleküle. Diese Signale regulieren die Proliferation und Differenzierung der sich entwickelnden Zellen.

In dieser Arbeit werden die Wechselwirkungen zwischen Blutzellvorläufern und ihrer Mikroumgebung aus verschiedenen Blickwinkeln und in verschiedenen biologischen Systemen untersucht. Zunächst entwickeln wir hierfür eine Methode zur Identifizierung von Liganden-Rezeptor -Wechselwirkungen auf der Grundlage transkriptomischer Daten. Mit dieser Methode können bekannte Wechselwirkungen wiedergefunden und neue Wechselwirkungen zwischen hämatopoetischen Stamm- und Vorläuferzellen und ihrer Nische im Knochenmark vorhergesagt werden. Unsere Analyse hebt insbesondere die einzigartige Rolle von mesenchymalen Stammzellen und Endothelzellen hervor und zeigt weiter, dass eine Nischenzellpopulation, die einen bestimmten Liganden exprimiert, häufig auch ihren Antagonisten exprimiert. Dieses Ergebnis deutet auf eine dynamische Regulierung von HSPCs durch ihre Nische hin.

Des weiteren quantifizieren wir in dieser Arbeit die Zellproliferation, ein Merkmal, das in vivo durch die Mikroumgebung reguliert werden kann. Wir entwickeln eine Methode, um die Länge jeder Zellzyklusphase sowie die relative GröSSe der ruhenden Zellfraktion zu bestimmen. Wir zeigen, dass unsere Methode, die aus der Kombination eines kurzen Puls-Chase-Experiments mit EdU, einem Thymidin-Analogon, und mathematischer Modellierung besteht, die genaue Bestimmung der Zellzykluslänge einer Krebszelllinie (Tet21N) ermöglicht. Wir wenden unsere Methode auf doppelt positive Thymozyten an und stellen fest, dass sich die meisten Zellen aus dieser Population im Ruhezustand befinden, ein kleiner Teil dieser Zellen jedoch mit einer mittleren Zellzyk-

luslänge von 9 Stunden, der Hälfte der Zellzykluslänge von Tet21N, stark proliferativ ist.

In einem letzten Teil der Arbeit untersuchen wir die T-Zell-Entwicklung im Thymus. Wir untersuchen, wie der murine Thymus, ein Organ, das normalerweise vom Zustrom von Knochenmarkvorläufern abhängt, autark werden kann, wenn er in eine Maus transplantiert wird, die keine kompetenten Knochenmarkvorläufer produziert. Wir zeigen, dass in diesem Rahmen, in dem thymuseigene Thymozyten nicht mit ankommenden Vorläufern um Il-7 konkurrieren müssen, eine bestimmte Population, die Doppeltnegative Drei (DN3), sich selbst erneuern kann, und das obwohl sie preTCR-Signale empfängt, die normalerweise ihre Differenzierung antreiben. Wir zeigen weiterhin, dass die längere Verweilzeit dieser DN3-Vorläufer ihre Differenzierung in γ δ T-Zellen ermöglicht, was normalerweise durch das preTCR-Signal ausgeschlossen ist. Unsere Ergebnisse geben Aufschluss darüber, wie die begrenzten Ressourcen, die im Thymus verfügbar sind, und der Fluss des ankommenden Vorläufers die zelluläre Antwort auf das preTCR-Signal beeinflussen und dadurch das Schicksal der DN3 direkt beeinflussen.

Contents

1 Introduction

1.1 From stroma to microenvironnment

By tradition, immunologists live in a simple world. In this world cells are divided in two categories, the immune cells which immunologists will know in great details, and all the other cells types which constitute the environment where the immune cell reside. The cells from this second category are regrouped under the designation "stromal cells". While immunologists have long considered the immune system independently of these stromal cells, it progressively became clear that these cells were critical to regulate immune cell functions within secondary lymphoid organs (reviewed by [Roozendaal and Mebius, 2011]) and their development in primary lymphoid organs.

The idea of a specialised cellular environnment dedicated to the regulation of an hematoietic cell population was first formulated by Raymond Schofield in 1978 [Schofield, 1978]. Schofield proposed that hematopoietic stem cell(HSC) quiescence and differentiation is controled by cells in physical contact with the HSCs, which he called niche, by analogy to the ecological niche (the set of environmental conditions required for maintenance of a living species). Schofield proposed that the niche ensures stem cell quiescence and therefore limits the accumulation of mutations in HSCs (as incured through division).

Since then, the existence of the hematopoietic niche has been documented extensively, it is now established that cells in the bone marrow provide HSPCs with cytokines regulating both proliferation and differentiation (reviewed by [Kumar and Geiger, 2017, Crane et al., 2017]). These parameters must be finely regulated to ensure self-renewal but avoid over-proliferation.

Some specific cytokines and their functions have been identified and studied intensively, such as Cxcl12, foremost homing cytokine, which maintains the HSPCs in close proximity to the niche cells ([Sugiyama et al., 2006]) and Stem Cell Factor (also known as Kit-ligand) which is critical to HSCs survival ([Ogawa et al., 1991]).

These Solubles factors contribute to HSCs regulation but cell surface adhesion molecule are also involved, they ensure stable contact between HSPCs and niche cells and it has also been proposed that this physiscal contact via cell adhesion molecule ensure proper balance between self renewal and differentiation via assymetric cell division ([Morrison and Kimble, 2006]).

The identification of the actual cell populations actively regulating HSPCs has proved difficult. Many cell types have been proposed over the years based on the effect on HSPCs of the specific ablation of these candidates: Osteoprogenitors, arteriole and sinusoidal endothelial cells megakaryocytes (reviewed by [Pinho and Frenette, 2019])and multiple populations of stromal cells identified by distinct markers (Osx [Rodda and McMahon, 2006], Nestin [Mendez-Ferrer et al., 2010], Prx-1 [Mizoguchi et al., 2014], Lepr [Zhou et al., 2014]).

It is not clear that all these cells identified by different markers are actually distinct. With that respect The team of our collaborator Claudia Waskow at the Technical University of Dresden has shown by cross correlating the expression of fluorescent lineage specific reporters and flow cytometry phenotyping that many of these populations of stomal cells actually overlap [Mende et al., 2019].

It is also unclear to what extent each distinct niche population is needed, what factors are expressed redundantly by several niche cell populations and what factors can only be provided by specific cell types. In other word, a global map of the interactions between HSPCs and their microenvironment is still needed if one wants to understand globally how niche cells can control HSPCs fates.

1.2 Quantification of cell dynamics

The microenvironment regulates the survival, proliferation and differentiation of the developing blood cells: in order to understand the effects of the environment, one needs tools to describe quantitatively these parameters in vivo.

Modern fate mapping technologies offer a way to quantify cell differentiation of blood cell progenitors in vivo without perturbing the biological system. In particular, our lab has been involved in estimation of cell differentiation from fate mapping data in partnership with the lab of Hans-Reimer Rodewald at the DKFZ. The lab of Hans-Reimer Rodewald has developed an inducible Cre-Lox system to label Tie2+ HSCs specifically. The aquisition of the label by the progeny of these intially labeled cells can then be followed over time. From the label equilibration dynamics between cell compartments, our lab was able to estimate rate parameters for differentiation ([Busch et al., 2015]). While proliferation rate of each of the cell compartment can be estimated from such data, it cannot resolve the actual cycling dynamics of the cell population which is here confounded with cell death and the potential existence of a quiescent fraction within the population.

Cell cycle parameters can now be quantified in vitro thanks to time-lapse micropscopy and the advent of cell phase specific fluorescent reporters such as the FUCCI markers[Sakaue-Sawano et al., 2008]. Recent studies have used such technologies to estimate both the duration of cell cycle phases and the quiescent fraction for cell populations in culture [Zerjatke et al., 2017, Tyson et al., 2012]. Intra vital imaging allows tracking of cells in vivo but cells cannot be tracked in this manner over several hours so that direct determination of cell cycle parameters in vivo remains an unresolved challenge.

A weapon of choice to measure cell proliferation in living organisms is the thymidine analogue incorporation assay. Generally the assay involves the treatment with a nucleotide analogue to thymidine (the only nucleotide unique to DNA) which can be detected specifically. This analogue is incorporated in DNA during replication and one can then detect the cells which have been through S phase while the analogue was present. The first techniques to detect nucleotides incorporation in DNA involved radiolabeled nucleotides which made the technique experimentally

cumbersome [Cavanagh et al., 2011]. A major advance in the field was the development of BrdU (Bromodeoxyuridine) detection assays in 1975 (reviewed by [Leif et al., 2004]). BrdU is a thymidine analogue which can be detected specifically with antibodies; it requires denaturation of the DNA to make the incorporated BrdU accessible to the antibody. Single cells which have incorporated BrdU can be identified by flow cytometry. More recently, alternative to BrdU have been used, most notably EdU (5-ethynyl-2'-deoxyuridine) [Salic and Mitchison, 2008]. EdU can be detected by Click-it chemistry reaction. The detection is more sensitive than BrdU, it does not require denaturation of DNA and requires less steps and time than BrdU detection. Since EdU detection does not involve DNA denaturation, it offers a better resolution of cell cycle phases by total DNA content labeling than BrdU.

BrdU pulse-chase experiments are sometimes used in combination with mathematical modeling to quantify cell proliferation [DeBoer and Perelson, 2013] but this proliferation rate is not to be confused with cell cycle rate, as it does not distinguish the cycling cells from the quiescent. Some authors have used progression of thymidine analogue in combination total DNA staining to estimate the duration of cell cycle phases [Vibert and Thomas-Vaslin, 2017, Kretschmer et al., 2020] but these papers also do not consider quiescent fractions and it remains unclear to what extent both cell cycle duration and quiescent fractions can be identified in vivo through this type of methods.

1.3 T cell development, a journey through the thymus

The Thymus is another biological system where cells fate is determined by the environment. The thymus was long considered a vestigial organ until Jacques Miller discovered it is a key organ of the immune system, the place where a specific subset of antigen specific immune cells, the lymphocytes T differentiate [Miller, 1961, Miller and Mitchell, 1967]. In the thymus, T cell progenitors recombine successively the β and α chains of the T Cell Receptor(TCR) to produce and are selected based on proper recombination and affinity of the TCR for Major Histocompatibility Complex+ peptide(MHC) presented by thymic epithelial cells and thymic dendritic cells. T cell receptor based selection is not however the only function of the thymus microenvironment. The contin-

uous production of a diverse T cell repertoire by the thymus depends on the strictly regulated, staged proliferation of the thymus progenitors which is coordinated with the recombination and TCR selection.

As the T cell progenitor progress down the differentiation path, their proliferation is regulated by different sets of cytokines produced by thymic epithelial cells (TEC). The early progenitors (DN1 and DN2) expension depends initially on Stem Cell Factor (SCF) ([Rodewald et al., 1995]) and, as they move to the DN3 stage, where the selection of the beta chain occurs, thymocytes lose the expression expression of the the SCF receptor and their survival and proliferation becomes dependent on Interleukin-7 and the Notch ligands Delta like and Jagged-1. Finally, once they reach the double positive stage, their differentiation, survival and proliferation depends on the cognate interaction with MHC expressing cells. This differentiation process is an actual journey through the thymus, the thymocytes migrate under the guidance of several chemokines, notably Cxcl2 and Ccl25 [Hu et al., 2015]. The Thymus is normally seeded continuously by progenitors from the bone marrow, but the nature and number of these progenitors remain highly discuted ([Krueger et al., 2017]). It has long been considered that a thymus could not function autonomously in a absence of competent progenitors. Indeed, experiments of transplantation of wild type thymi into recipient mice lacking Rag2, a key enzyme in TCR recombination, have shown rapid depletion of the donor thymus ([Frey et al., 1992]), by lack of competent progenitors. In 2012 however, the groups of HR Rodewald at the DKFZ, and Benedicta Rocha at Institut Pasteur in Paris discovered in paralell that, under some condition T cell development can be maintained in absence of progenitors [Martins et al., 2012, Peaudecerf et al., 2012]. Indeed Rag2 Il2rγ chain (component of the Il-7 receptor) KO mice transplanted with wild type thymi can keep producing T cell for several month while resident thymocytes are eliminated over a month when a thymus is transplanted into recipient which produces competent Bone marrow progenitors. Of note, the thymus autonomy not a stable situation as a T-cell lymphoblastic leukemia will develop in the thymus graft around 16 weeks after transplantation in a majority of cases [Martins et al., 2014]. The authors concluded from these experiments that competition for Il-7 is a key factor driving the replacement of resident thymocytes by incoming bone marrow progenitors. And that in absence of progenitors able to compete for Il-7, the resident thymocytes can self sustain. While it was proposed that this competition happens early in T cell development The precise stage of T

cell development where the competition for Il-7 occurs and the actual changes in cell dynamics involved remain to be precised.

Il-7 is critical to T-cells and their progenitors. In the periphery, it is involved in homeostatic proliferation of naive and memory T cells [Surh and Sprent, 2008] and in the thymus, it plays multiple role in the course of thymocyte development. In the early DN2 stage, Il-7 sends a pro-survival signal by inducing the expression of the antiapoptotic protein Bcl-2. The role of Il-7 receptor at the γ selection stage remains under discussion, [Xiong et al., 2013] Xiong and colleagues reported that Il-7 plays the same role as in DN2 while Boudil et al. [Boudil et al., 2015] proposed that the cytokine induces further transcriptional changes independent of Bcl-2 although the precise functional effects of these changes remain unclear. Furthmore, distinct papers have reported with in vitro assay that Il-7 signaling blocks differentiation at the DN3b stage [Tussiwand et al., 2011, Boudil et al., 2015]. At the DN4 stage, promotes cell proliferation and delay TCR α rearrangement. the expression of the Il7rα is then lost and Il-7 is not expected to play a role downstream. Interestingly, a dose effect of I-7 on T cell development has been reported [El Kassar et al., 2004] in mouse model overexpressing the cytokine but it remains unclear whether it plays a role in the various effect generally observed in normal conditions.

The thymus autonomy model could offer an opportunity to understand how proliferation and differentiation are influenced by the thymic microenvironment at this critical stage of T cell development.

1.4 Objectives

In this work, we study interactions between blood progenitors and their niche through multiple angles and methods. First of all, we seek to characterize the molecular actors of HSPCS regulation by bone marrow niche cell at steady state. Using population level RNA sequencing data produced in the lab of Claudia Waskow we intend to infer potential interactions between niche cells and HSPCS and to obtain a map of HSPCS-niche interactions. With this approach, we will not get a dynamical picture of stem cell interactions with the niche but the expression patterns of these interactors can nonetheless provide some hints regarding the ability of the niche to regulate stem cell function in a dynamic fashion.

In a second part, our aim is to develop a tool to quantify cell cycle parameters in vivo based on EdU labeling. Cell proliferation is one of the mode of action of the niche on developing cells and such a tool can prove useful in the characterization of the action of the niche.

The model of thymus autonomy developed in the lab of Hans-reimer Rodewald constitutes an invaluable tool study the regulation of the differentiation of early thymocytes. In a third part, we intend to use of thymus autonomy developed in the Rodewald lab to study the role of the thymic niche, and in particular Il-7 in the development of early thymocytes.

2 Material and Method

2.1 Material

Lab equipment

equipment	**Manufacturer**
Cellometer Auto 2000	Nexcelom
Digital thermometer Qtemp 200	VWR International
Heraeus Fresco 17 Centrifuge	Thermo Fisher Scientific
Heraeus Megafuge 40R	Thermo Fisher Scientific
LD79 Digital Test-Tube Rotator	Labinco
LSRFortessa (Flow Cytometer)	Becton Dickinson
FACSAria III (Cell sorter)	Becton Dickinson
Thermomixer compact	Eppendorf
MACSQuant Analyzer 10 (Flow Cytometer)	MACS
Steri-cult CO2 incubator	Thermo Fisher Scientific
Cell culture hood HERA	Thermo Fisher Scientific

Consumables

Description	**Manufacturer**
Cell strainer (40 µm)	Falcon
Cellometer Disposable Imaging Chambers	Nexcelom
Eppendorf safe lock tubes (0.5 ml, 1.5 ml, 2 ml)	Eppendorf
Falcon tube (15 ml, 50 ml)	Falcon
Injekt Syringes (2 ml)	Braun
Multidish 6 well, 96 well concial bottom plates	Nunc
Sterican needles (26 G)	Braun

Antibodies

specificity	clone	Fluorochrome	dilution	manufacturer
CD4	H129.19	PE	1:200	BD Pharmingen
CD8a	53-6.7	APC	1:100	eBioscience
CD8a	53-6.7	PE	1:200	BD Pharmingen
CD25	PC61	PeCy7	1:600	BD Pharmingen
CD44	IM7	PercpC5.5	1:100	BD Pharmingen
TCRb	H%/-597	APCeFl780	1:50	eBioscience
TCRdelta	GL3	FITC	1:200	eBioscience
TCRdelta	GL3	APC	1:50	eBioscience
CD117	ACK-2	BV650	1:200	BioLegend
CD117	ACK-2	APC Cy7	1:400	eBioscience
CD3	17A2	APC eFl780	1:25	eBioscience
CD45.1	A20	A700	1:400	eBioscience
CD45.2	104	FITC	1:50	BD Pharmingen
CD19	6D5	BV421	1:400	BioLegend
CD11b	M1/70	BV421	1:800	BioLegend
Ter119	PER-119	BV421	1:200	BD Pharminogen
NK1.1	PK136	BV421	1:100	biolegend
TCR gd	GL3	BV421	1:100	BioLegend
GR-1	RB6-8C5	BV421	1:800	BioLegend

2.2 Method

2.2.1 Animal Procedures

For every EdU experiments we used B6.129X1- Gt(ROSA)26Sortm1(EYFP)Cos/J mice available in the Rodewald lab. These mice derive from C57BL/6 mice; they carry a loxP-flanked STOP sequence followed by the Enhanced Yellow Fluorescent Protein gene (EYFP) inserted into the Gt(ROSA)26Sor locus. EYFP can only expressed following deletion of the STOP sequence by a CRE recombinase. For the thymus transplantation experiments, recipient mice were C57BL/6, Rag2$^{-/-}$ Il2r $\gamma^{-/-}$and donor thymi were B6/Ly5.1 (H-2^{b}; CD45.1^{+}) Mice were kept in individually ventilated cages (IVC) under specific pathogen free conditions in the animal facility of the German Cancer Research Center (DKFZ) and provided with autoclaved drinking water and food ad libitum. All animal procedures were performed according to national and institutional guidelines (DKFZ Heidelberg) and approved by the Regierungsprasidium Karlsruhe.

2.2.2 thymus transplantation

Thymus transplantation under the Kidney capsules were performed by Csilla Kongsaysak-Lengyel. Details on the transplantation are available in her doctoral thesis (University of Heidelberg)

2.2.3 Preparation of murine thymocytes

Mice were killed by CO2 inhalation. Thymi were smashed through a 40 µm filter with the plunger of a 5 ml syringe and resuspended in 5ml of FACS buffer. cells were counted with a Cellometer 2000 cell viability counter according to the manufacturer s instructions.

2.2.4 Immunostaining

Every further steps of cell surface Immunostaining was performed in FACS buffer (PBS+5% FCS) at 4°C. To prevent non specific antibody binding to Fc receptor, cells were incubated 20 minutes with 0.24 mg total mouse igG. Next an antibody mixed was added to the solution (final volume 300 µl for 10^7). Cells were incubated for 30 to 45 minutes in the staining mix before being was washed with either FACS buffer or PBS+5% BSA for further processing with the EdU detection kit(cf. 2.2.5).

2.2.5 EdU assay on thymocytes

Click-iT Plus EdU Alexa Fluor 488 flow cytometry kit (Invitrogen) was used for EdU labeling and detection.

EdU (Invitrogen) was diluted in sterile PBS to a final concentration of 2.5 mg/ml. Aliquots were stored at -20°C until usage. Mice were injected i.p. with 1 mg EdU diluted in 400 µl PBS.

1h, 2h, 3h, 4h, 5h, 8, 12h, 14h and 18h after injection, mice were sacrified, and thymi collected and counted according to 2.2.3 and 10 to 15 millions thymocytes were then resuspended in 100 µl Zombie Red solution (1:1000 in PBS) for 15 minutes at room temperature in order to detect dead cells. Cells were then stained with antibodies according to 2.2.4. Thymocytes were then fixated, permeabilized using the permeabilization buffer from the EdU kit (15 minutes incubation at room

temperature), the Click-it reaction cocktail was then added to the solution. After 30 minutes of incubation at room temperature, cells were washed in Click-it permeabilization buffer. For intracellular staining was performed, cells were further blocked for aspecific binding sites with total igG (see 2.2.3) diluted in Click-it Permeabilizing buffer for 20 minutes and then antibody solution (antibodies diluted in click-it perm buffer) was added and the cells were incubated for 30 minutes. cells were then washed and resuspended in Click-it Permeabilization Buffer. 30 minutes prior to FACS analysis cells were resuspended in 1 ml/10^7 FxCycle Violet solution (1:1000 in Click-iT permeabilization buffer) . Cells were analyzed on a BD LSRFortessa. For correction of spectral overlaps, single stained compensation beads were measured for each fluorochrome before each series of measurement.

2.2.6 Cell sorting

For the purpose of the single cell RNA-seq experiment, cells were sorted on a BD LSRFortessa cell sorter, into PBS+10 % FCS. After sorting, cells were centrifuged at 300 g for 5 mins and resuspended into 35 µl PBS+10% FCS, they were then maintained at 4°C to be further processed according to 10xgenomics Chromium Single Cell 5 Library Construction Kit and Chromium Single Cell V(D)J Enrichment Kit, Mouse T Cell.

2.2.7 Tet21N EdU assay

Human neuroblastoma derived Tet21N cells were grown in RPMI 1640 medium supplemented with 10% fetal calf serum at 37°C, 5 % CO2and 88% humidity.

$1.5 \cdot 10^6$ cells were seeded on 15 cm Petri dishes one day before EdU treatment and then treated with EdU at a final concentration of 10 µl Plates were washed twice with medium one hour after incubation to remove EdU in the medium. After 1h , 3h, 4h, 5h, 8h, 12h, 14h, 18h of incubation Cells were detached from the dish by incubating them with 5 ml Versene 5 minutes at 37 C, they were counted and 1 M cells per sample were transfered in 15 ml FALCONS and fixed by incubating the cells in 250 ul PFA 15 minutes at room temperature in the dark. The fixation reaction was stoppped by adding 15 ml PBS. Cells were then resuspended in 300 µl and 3ml ice cold methanol

was added slowly drop by drop. Cells were then kept at -20°C for at least one day.

On the day of the FACS analysis, cells were washed by first adding 10 ml PBS to the tubes and centrifugating 10 minutes at 2500 RPM. Cells were then resuspended in 4 ml washing solution (for 500 ml washing buffer: 475 ml sterile water 25 ml PBS 20X, 5g BSA, 500 µl Triton-X), and incubated for 30 minutes to rehydrate the cells. Cells were then resuspended in PBS+1% BSA, then cells were permeabilized and Click-it EdU reaction was performed according to the manufacturer's instructions.

The cells were finally resuspended in 200 mul in Fx Cycle solution (1:1000 in Click-it Permeabilization buffer)

2.3 Bioinformatic analysis

2.3.1 Bulk RNA-seq analysis

The analysis of the niche cells and HSPCS RNA-seq data produced by the Waskow lab in Dresden was performed with R software v3.3.2. Read count was performed using GenomicAlignments package v3.2.3, with mouse annotation GRCm38.81 and read counts were normalized with the regularized log2 transformation (rlog) implemented in the R package Deseq2. Transcriptome data sets were corrected for batch effects and principal component analysis (PCA) was computed using FactoMineR package v 3.3.2. the batch effect between niche cell samples was corrected with the ComBat algorithm implemented in the R package SVA. Hierarchical clusterings based on Euclidian distance were performed with the hclust function implemented in R. Heatmap was generated with the rlog-transformed read counts using the Pheatmap package of R. Networks displayed were created with cytoscape v3.4.0 software [Shannon et al., 2003].

2.3.2 Single-cell RNA-seq analysis

Single cell RNA library of DN3 thymocytes were generated by Nina Claudino with the 10X genomics Chromium Single Cell 5' Library and Gel Bead Kit and Chromium Single Cell V(D)J En-

richment Kit, Mouse T Cell. And the libraries were sequenced at the DKFZ sequencing core facility. the single cell data was then aligned to the mouse genome (mm10) and reads per gene were counted with the softare Cell Ranger v 2.1 Row count matrix were then checked for quality with R according to the Bioconductor pipeline. Cells were rejected based on low UMI counts, low number of expressed genes and high proportion of Mitochondrial genes with the isOutlier function of the R package Scater. Cell size factors were estimated using the calculateSumFactors function of the package Scran and finally counts were log2-normalized using the function normalize of Scater,

Sequences from the T cell receptor enriched library assembled by Cell Ranger were realigned to the mm10 mouse the genome using Blastn and contigs containing sequences comprised between the D1 and J1.1 or D2 and J2.1 regions of the TCR beta genes were categorized as "germline" as these regions are removed during the D to J joining.

2.3.3 Estimation of model parameters

Data assignement to the model

Populations of interest are gated based on total DNA content and EdU label. Based on total DNA content, we can define a G0G1 gate, an S gate, and and a G2M gate. In every biological systems that we that we study, the distribution of cells in these gates is considered constant. We also collect the fraction of EdU labeled cells in these gates over time which are initially labeled in S phase and then progress to the other cell cycle phases. Our dataset therefore comprises the steady state values of the S and G2M gates (these 2 values are sufficient to describe the full cell cycle distribution since the total compartment is defined as G0G1+G1+G2M) as well as the fraction of EdU labeled cells in the G0G1, S and G2M gates at each time point. DNA does not allow perfect discrimination of cell cycle phases, it is known that a certain fraction of cells in S phase in comprised within the G0G1 and G2M gates. When fitting the model to the data we therefore assign a certain number of subphase of the model denoted m_{G0G1} and m_{G2} for the G0G1 and G2M gates respectively (Fig. 7 D).

We used Monte Carlo Markov Chain to establish the credible interval of the parameters given the data.

Credible interval estimation with Monte Carlo Markov Chain

The Monte Carlo Markov chain is a method that generates iteratively samples from a probability distribution until the target distribution is approximated. At each iteration a new set of parameter is generated and can be either accepted or rejected based on its probability relative to the current set of samples.

In our case, we seek to approximate $P(\theta|D)$ the probability distribution of our parameters θ given the data D. Pursuant to Bayes Theorem, we can write:

$$P(\theta|D) = \frac{\mathscr{L}(D|\theta)P(\theta)}{P(D)} \tag{2.1}$$

$\mathscr{L}(D|\theta)$ is the likelihood function. $P(\theta)$, called prior distribution, encodes prior knowledge about the parameters, in our case, we assume uniform prior distribution within reasonable bounds set by our biological system and we can therefore ignore this term. $P(D)$ is the total probability that such data arises, this normalising factor is also irrelevant to our parameter estimation.

Consequently we can write :

$P(\theta|D) = \mathscr{L}(D|\theta)$

For any given θ, we compute the likelihood function in a standard way:

$$\mathscr{L} = \sum (\frac{O_i - M_i}{S_i})^2 \tag{2.2}$$

Where O_i is the mean observed value for a given data point, M_i is the relative value predicted by the model at the same time point, and S_i is the standard error to the mean value.

The Monte Carlo Markov chain was performed using Emcee [Foreman-Mackey et al., 2013], an algorithm implemented in the Python programming language. Emcee is an Ensemble sampler, it runs several Monte Carlo Markov chain in parallel. At each iteration, the current value of each chain is updated based on the current values of the other chains. In accordance with the programme's guidelines, we considered a chain to have reached steady state once it was 50 times longer than the autocorrelation times of parameter.

3 Results

3.1 Inference of molecular interactions between HSPCs and niche cells in the mouse bone marrow

Several bone marrow cell population have been described as contributing to HSPCS maintenance [Pinho and Frenette, 2019]. The respective contribution and importance of each of these populations to HSPCS maintenance is still however unclear. It is also unclear whether each of the separate populations described by different research groups as identified by distinct cell surface marker actually transmit distinct signals to HSPCs and are therefore indispensable.

It has been shown that HSPCS are not distributed uniformly in the bone marrow. Some authors have suggested that the most quiescent HSCs tend to localize preferentially on the endosteal region of the bone, near the hard bone [Zhang et al., 2003, Kunisaki et al., 2013] while other studies have stressed that HSCs from the central bone marrow have distinct properties from their endosteal counterparts and tend to be more proliferative but with lower homing and reconstitution potential [Grassinger et al., 2010].

In order to test whether the function of the niche cells generally depends on their localization in the bone marrow, and in particular, whether cells from the central bone marrow generally have distinct properties from the cells localized in the endosteum, the group of Claudia Waskow at the Technical University Dresden isolated separately cells from these two compartments. In each of these two compartments, they isolated majors components of the bone marrow environment which had been described as important support of hematopoiesis namely $CD31^{-}CD144^{-}CD51^{+}$ $Pdgfra^{+}Sca1^{-}$ Mesenchymal Stromal Cells (MSC) and $Pdgfra^{+}Sca1^{+}$S

PaS cells [Morikawa et al., 2009], CD31^{hi} CD144^{+} arterioral enthotelial cells (Endo) [Ding et al., 2012], and CD31^{-}CD51^{+} Pdgfra^{-} osteoprogenitors (OP) [Taichman et al., 1996, Calvi et al., 2003], (Fig. 1).

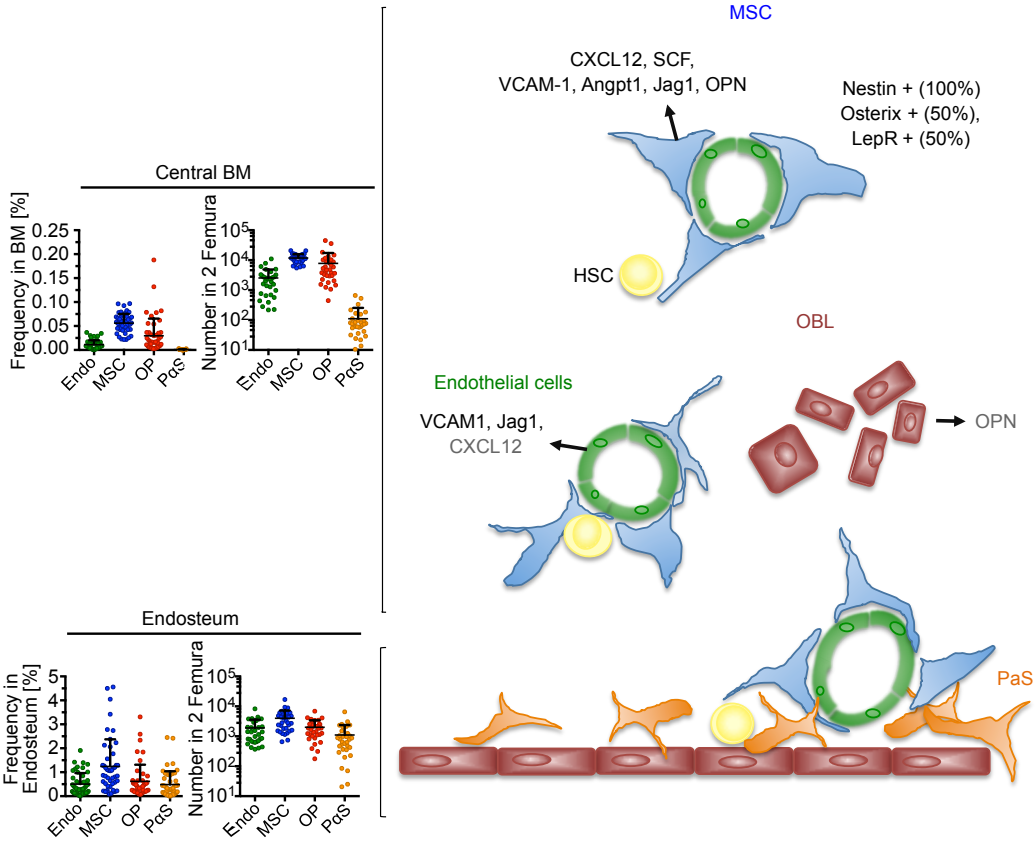


Figure 1: Bone marrow niche populations isolated by Nicole Mende with their respective frequencies in the mouse central bone marrow and Endostheum.

Population level RNA-seq was performed on each of these niche populations as well as on the population at the top of the hematopoietic hierarchy: long term HSCs (LT-HSC; Lin^{-}Kit^{+}Sca1^{+}CD48^{-}CD150^{-}), short term HSCs(ST-HSC; Lin^{-}Kit^{+}Sca1^{+}CD34^{+}CD135^{-}) and MultiPotent Progenitors (MPP;Lin^{-}Kit^{+}Sca1^{+}CD34^{+}CD135^{+}) as defined by [Yang et al., 2005]. We used the gene expression data to infer candidate interactions between these HSPCS and each niche cell type, and determine to what extent, each niche cell type is providing distinct ligands to the HSPCS.

3.1.1 Ligand expression distinguishes niche cell types

database of receptors ligand interactions

We first established a database of ligand-receptor pairs. Most protein-protein interactions(PPI) databases such as Intact [Orchard et al., 2014], Biogrid [Stark et al., 2006], or MINT [Chatr-aryamontri et al., 2007] contain mainly interactions between human genes. As of the 7th of January 2020 Intact DB lists 611,461 binary interactions between human proteins and 84,042 interactions between mouse proteins. We decided to retain interactions between human proteins with orthologues in mouse. Such databases of receptors-ligands have been produced over the years, of which the Fantom V database [Ramilowski et al., 2015], with 1849 experimentally verified interactions between human genes, was the largest at the time this was work was performed. We used the Fantom V database as a basis and expended this database by including additional protein-protein interactions from several public PPI database which had not been searched by the authors of the Fantom V database, namely IntactDB, EBI-GOA non Intact and InnateDB [Lynn et al., 2008]. We selected interactions between genes annotated with the Gene Ontology(GO) term cell surface(receptors) and genes with GO annotation receptor binding (ligands). We further curated these interactions by searching the literature; as a result 370 interactions were added to the Fantom V database. We also excluded from the Fantom V database some substrate-enzyme type interactions which can occur at the surface of cells but whose role in cell to cell communication remains uncertain; 400 interactions were excluded based on this criterion. We ended up with a database containing 1863 interactions involving 555 unique receptors and 579 unique ligands.

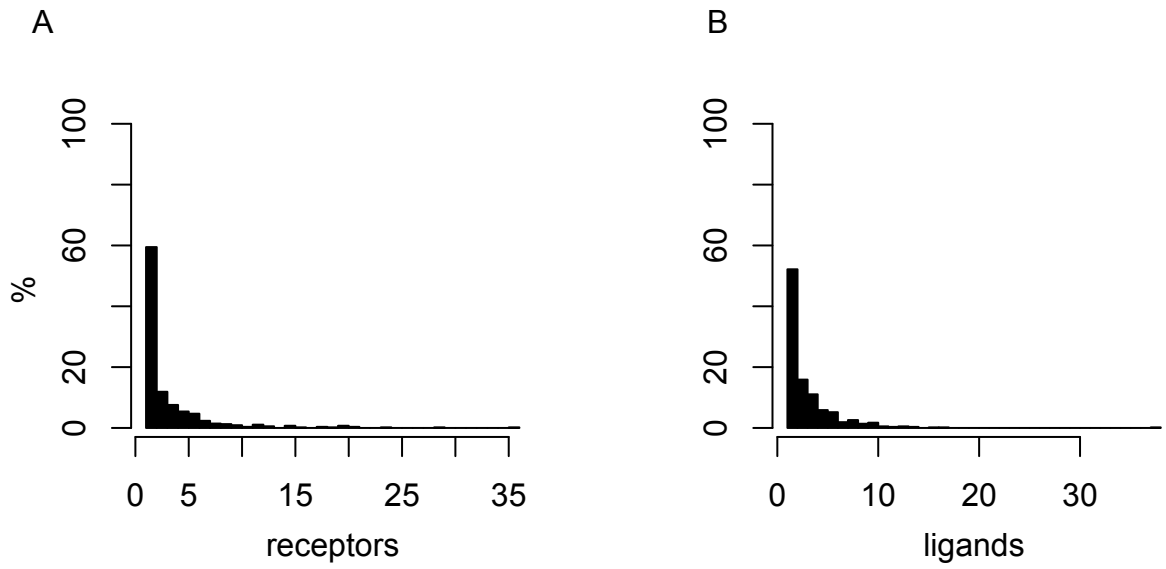


Figure 2: A. distribution of the number of ligands per unique receptor, B. number of receptors per unique ligands

Obviously, the relatively low number of unique interactors compared to the total number of interactions means that some interactors are involved in multiple interaction. The majority of interactors appear only once (Fig. 2) and more than 95% of receptors and 97 % of ligands appear nor more than 10 times. These receptors with large number notably include integrins, which are known to bind a great variety of extra cellular matrix protein and cell surface molecules [Humphries et al., 2006].

Eache niche cell type presents a distinct ligands expression profile

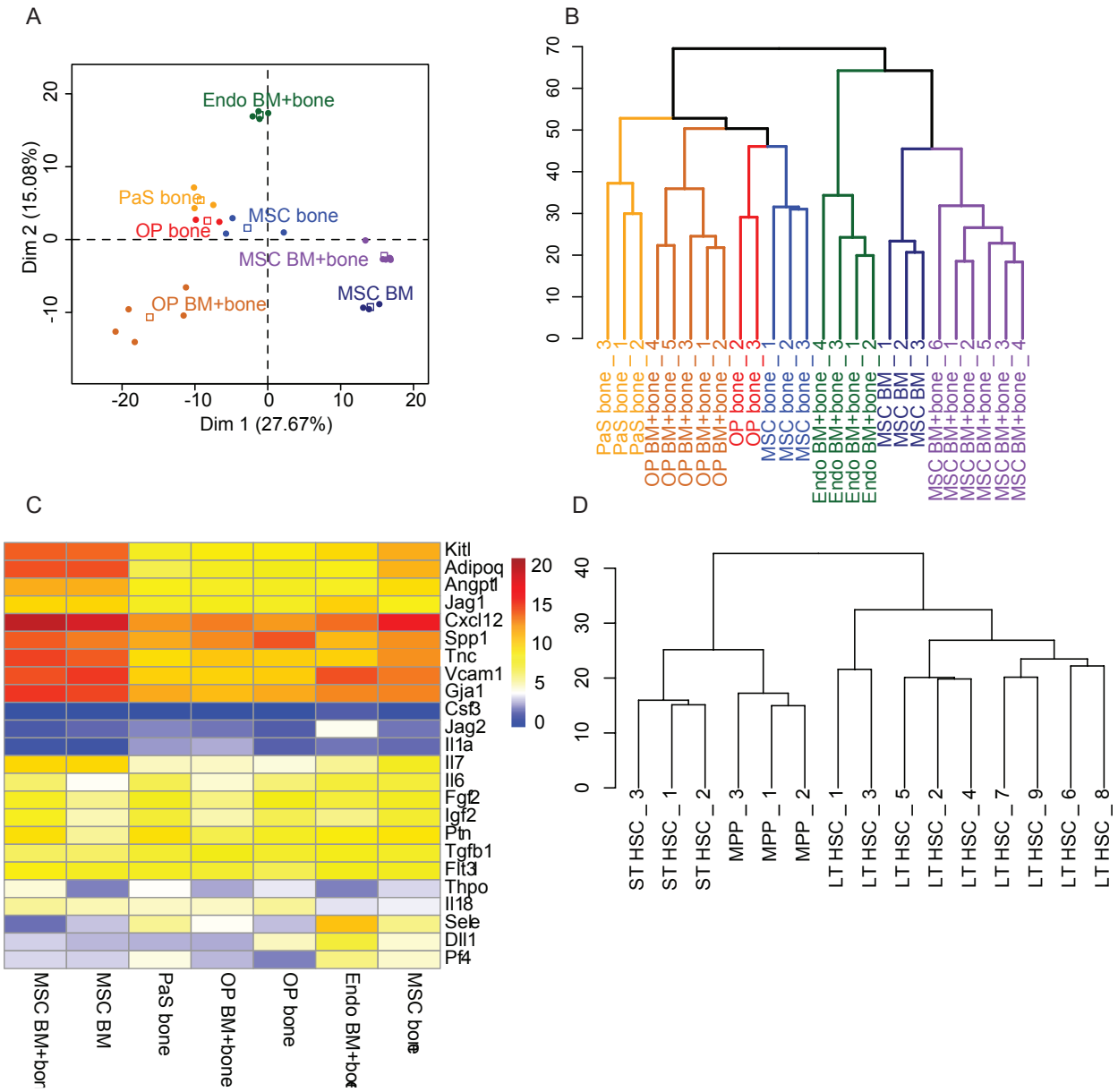


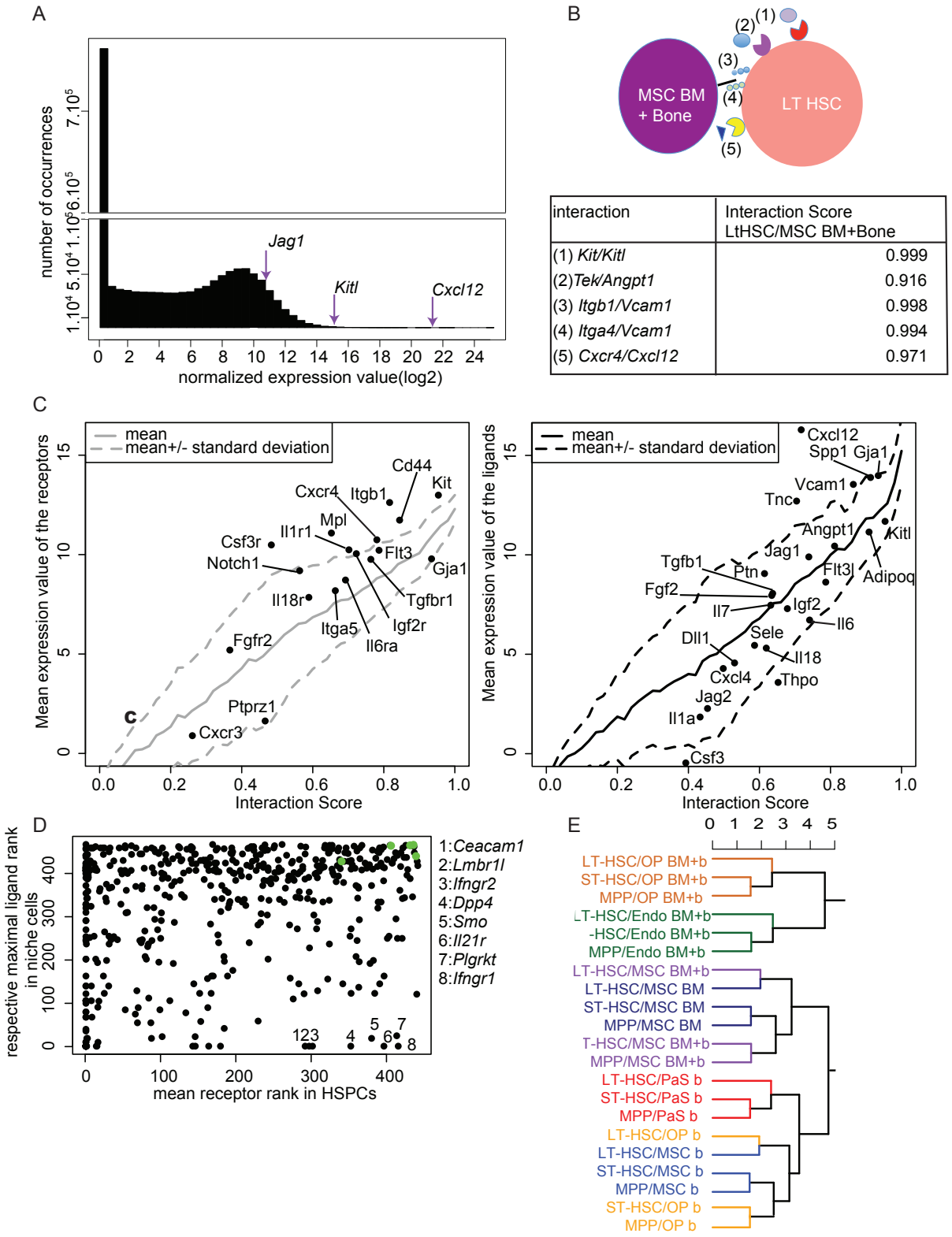
Figure 3: Analysis of ligand expression in the niche populations. (A) PCA based on gene expression of the 579 ligands from the receptor-ligand database by niche cells. Dots, individual samples; open squares, barycenters of the different samples for a given cell population. (B) Hierarchical clustering of niche cell samples based on ligand expression. (C) Heatmap shows mean rlog transformed expression values of known HSPC regulators (ligands) in distinct niche cells.(D) Hierarchical clustering of HSPCs based on gene expression of the 555 receptors from the receptor-ligand database.

To understand how HSPCS are regulated by niche cells we analysed the expression by HSPCs of receptors in the database and of the ligands by the niche cells. A Principal component analysis was

performed on the niche cells (Fig. 3 A). We can observe that the first component separates MSC BM and MSC BM+B from the other populations. It is known that MSC of the central bone marrow represent 80% of the MSCs; MSC BM+B are therefore MSC BM in majority. we can therefore conclude that the first component reveals a ligand expression pattern which is specific to MSC BM. Likewise, the second component separates the endothelial cells from the rest. Finally, the three cells populations isolated from the bone are not separated by the third first components, this observation suggests that niche populations isolated from the bone all show relatively similar ligands expression pattern. This observation is strengthened by hierarchical clustering showing that these niche cells of bone origin cluster together on a hierarchical clustering tree 3B. Figure 3 C shows the gene expression of some ligands which have been reported as participants in HSPCs regulation. We observe generally high expression of the key cytokines Cxcl12, involved in homing of the Hspcs in the bone marrow, and Kit, also known as stem cell factor, the key cytokine for survival of HSCs. The expression of these cytokines is particularly high in MSCs, a population which has already been described as an important provider of Cxcl12 [Morikawa et al., 2009].

3.1.2 Scoring the interactions

Ligands known to be expressed by specific niche cell types vary considerably in gene expression level. In fact, we observed the expression of Cxcl12 in MSC BM+B is 7.5 fold a high as the expression of Jag1 in endothelial cells (Fig. 3 A). This being said, compared to the overall distribution of expression values within our dataset, both these cytokines can be deemed highly expressed. This suggests that the rank of a particular expression value within a dataset contains information on the functional importance of this gene. A method based on rank preserves the information while reducing the empasis on extreme values. Moreover, rank-based comparisons can facilitate the integration of expression date from different experiments and platforms. We therefore decided to analyse the ranks of the interactions rather than the absolute expression values themselves and we developed a method to score interactions between niche and HSPCs based on the ranks of the receptor and ligands in our database.



interaction	Interaction Score LtHSC/MSC BM+Bone
(1) *Kit/Kitl*	0.999
(2)*Tek/Angpt1*	0.916
(3) *Itgb1/Vcam1*	0.998
(4) *Itga4/Vcam1*	0.994
(5) *Cxcr4/Cxcl12*	0.971

Figure 4: Scoring the interactions between HSPCs and niche cell populations.(A) Distribution of expression Values in the niche dataset. Arrows indicate expression of Kitl, Cxcl12 and Jag1 in MSC BM (B) Example of interactions and respective interaction scores for the LT-HSC/MSC (BM plus bone) pair. (C) Left panel, Relationship between mean expression of receptors in HSPCs and scores of interactions involving indicated ligands. Right panel, Relationship between mean expression of ligands in niche cells and scores of interactions involving indicated receptors. (D) Mean receptor rank within HSPCs (x-axis) vs the highest rank of the respective ligands in any niche cell (y-axis), green dots are, from left to right, Tek, Cxcr4, Itga4, Itgb1, Kit. (E) Hierarchical clustering of HSPC/niche cell pairs based on interaction scores

We first rank the receptors and ligands which are part of our interaction database. One ranking is conducted for each cell type, the ligands are ranked based on the mean expression values of the ligands in the niche cells, and the receptors are ranked based on the mean expression values of the receptors in the HSPCs.

For each interaction I, each HSPC H and niche N, a score S is computed as the sum of the rank of the receptor R and the ligand L.

$$S_{HNI} = rank(L_{NI}) + rank(R_{HI}) \tag{3.1}$$

This score is then normalized to the highest Score of the specific Niche-HSC pair NH. This interaction score IS has consequently a 0-1 range.

$$IS_{NHI} = \frac{S_{NHI}}{max(S_{HN})} \tag{3.2}$$

Validation of the method

In order to confirm the ability of our method to isolate important interactions, we selected 5 extensively studied interactions between MSC and HSCs (Fig.4 B) namely Kit ligand (Kitl) with Kit,Tek (Tie2) with Angpt1, Cxcl12 with Cxcr4 , as well the VLA-4 (Itgb1[integrin b1]/Itga4 [Integrin a4]) integrin complex with Vcam1. All these interactions appeared within the top 3% of the MSC/LT-HSC interactions.

We further observed (Fig.4 C)) in the interaction score correlated very well with both the expression of the ligand and of the corresponding receptor, although the ligands had overall higher expression values. Thus, the interaction score preserves the information for both ligands and receptors.

We also confirmed that our method worked in other systems. In particular, we selected expression data from mouse T cells and Dentritic cells from the Immgen database, a public database

of high quality trancriptomic data from murine immune cells [Heng et al., 2008]. The top 10 interactions found according to our method were the most described interactions between antigen presenting cells and T cells (Sup. Fig. 1). Altogether, these results support the use of our method as a means to study cell-to cell interactions.

Other candidate niche cells have been described besides the populations selected by our collaborators, including other hematopoietic cells: megakaryocytes and macrophages. It was therefore useful to check to what extend the niche cells that had been selected by the Waskow lab could provide ligands to the receptors expressed by HSCs. For the vast majority of receptors expressed in HSPCs (high rank) at least one cognate ligand was expressed by one or several niche cell population(Fig. 4 D). However, there were notable exceptions (numbered from 1 to 8 in Figure 4 D), including receptors for the proinflammatory cytokines interferon gamma and Il-21. Both of these cytokines are produced by activated T cells and NK cells and can be involved in HSCs response to stress [Parrish-Novak et al., 2002, Morales-Mantilla and King, 2018] but are less likely to play a major role in steady state adult hematopoiesis. These observations validate our niche cells population as major providers of ligands to HSPCs at steady state.

Hierarchical clustering of the niche/HSPC pairs based on the Euclidian distance between interaction scores showed a clear ordering according to the niche component (Fig. 4 E). Within each major cluster dictated by niche cell type, the interactions formed by LT-HSCs clustered apart from the interactions formed by both ST-HSCs and MPPs in all cases. These findings suggest that each niche cell type presents a distinct quantitative ligand profile to the HSPCs, while secondarily, the interactomes also show some difference between LT-HSCs on one side versus ST-HSC and MPPs on the other side.

Fig.4 E) also reveals that pairs involving niche cell isolated from the bone cluster appart from the rest, suggesting that distinct niche cell populations from the endosteum express the same ligands which set these populations apart from the niche cells of the central bone marrow. This hints at a regional effect of HSC regulation by the niche.

3.1.3 Niche cell-specific interactions

Specificity score

Following this finding, we aimed to determine which interactions are specific to a particular niche cell type. For this purpose, we then computed, for each interaction I the mean Interaction Score μ_I and its respective standard deviation σ_I

$$\mu_I = \frac{1}{n \times h} \sum_{i=1}^{n \times h} IS_{Ii} \tag{3.3}$$

$$\sigma_I = \sqrt{\frac{1}{n \times h} \sum_{i=1}^{n \times h} \left(IS_{Ii} - \mu_I\right)^2} \tag{3.4}$$

where h is the total number of HSPC types and n is the total number of niche cell types.

We then calculate for each niche cell N, a mean Interaction Score μ_{IN}.

$$\mu_{IN} = \frac{1}{h} \sum_{i=1}^{h} IC_{INi} \tag{3.5}$$

And we finally computed the deviation of μ_{IN} to μ_I.

$$Z\text{-}Score_{IN} = \frac{\mu_{IN} - \mu_I}{\sigma_I} \tag{3.6}$$

For each interaction, this Z-score measures the distance of the interaction score associated with a particular niche population to the mean interaction score. The highest the Z-score, the more specific it is to the niche population.

Specific ligands presented by endothelial cells and MSCs

We focused on endothelial cells and MSCs, which are often seen a the most critical providers of ligands to HSCs [Pinho et al., 2018]. We selected interactions that involve ligands with particularly high expression in one of these two niche cell types (Figure 5). Among these top niche cell-specific predicted interactions, we found several previously described ones, such as Cxcr4/Cxcl12 and Kit/Kitl associated with MSC, or Notch/Jag1, Cd44/Selectin E or Selectin L/Selectin E associated with endothelial cells. Furthermore, the two niche cells expressed distinct components of the extracellular matrix that can interact with the same cell adhesion molecule on HSPCs, such as collagen (endothelial cells) and laminin b1 (MSCs) interacting with integrin b1 (HSPCs).

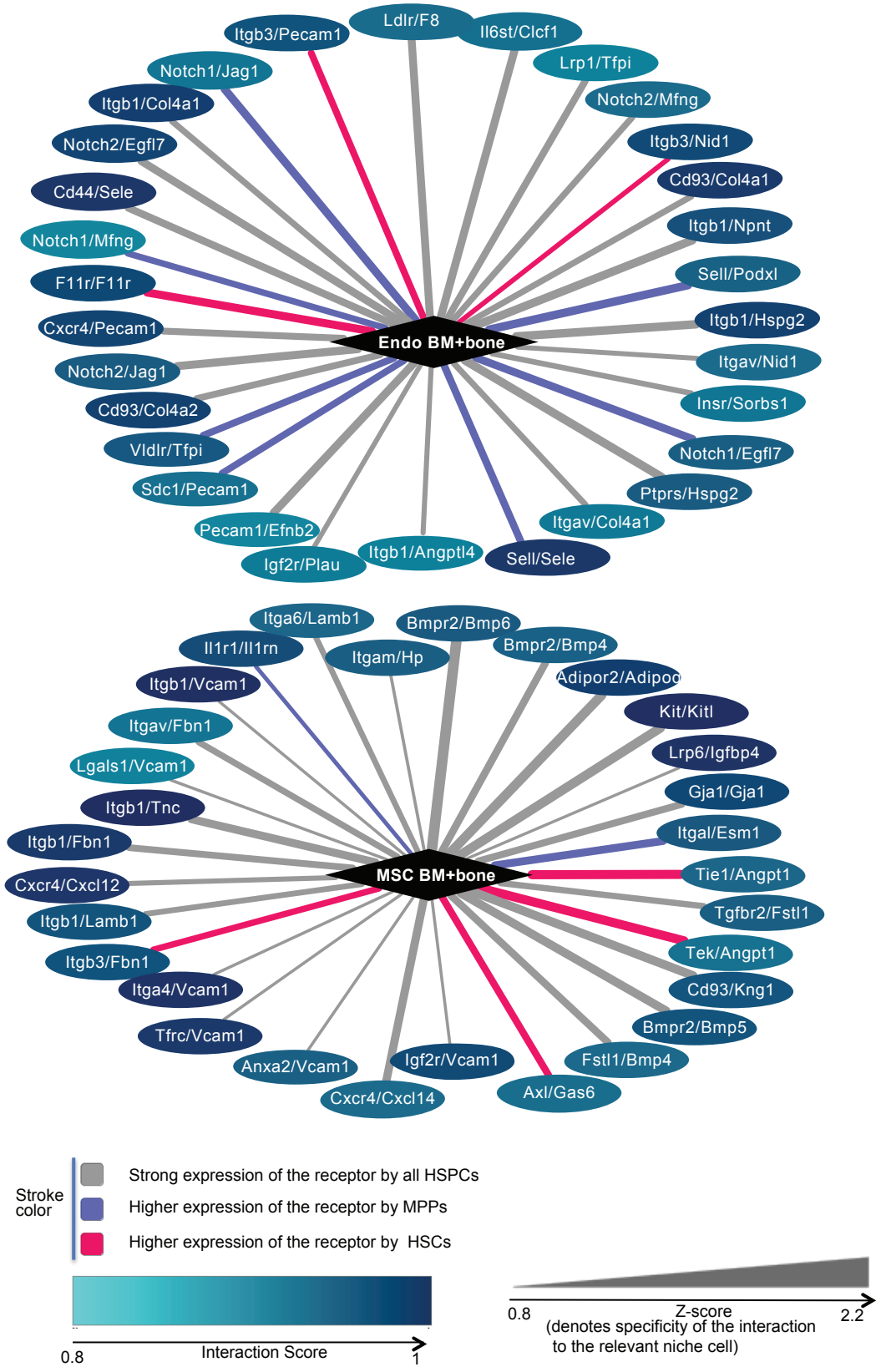


Figure 5: Receptor/ ligand interactions most specifically associated with endothelial cells BM+bone (upper network) and MSCs BM+bone (lower network). Node color encodes the interaction score of the specified niche cell with the type of HSPC for which the score is highest. Edge color visualized whether the highest interaction score is with LT-HSCs (red), with MPPs (blue), or whether the score is approximately equal for all HSPCs (gray). Edge thickness denotes the specificity of the interaction score to the niche cell type (Z-score).

identification of the source of known HSPCs ligands

Our analysis was also informative on receptor-ligand interactions which have been shown to participate in HSPC homeostasis but for which the cellular source of the ligand is not well defined. For instance, Axl/Gas6 appears as one of our top interactions with Gas6 being presented specifically by MSCs. Gas6 is known to maintain hematopoiesis in vivo and to support HSCs in culture via th Axl [Axelrod and Pienta, 2014], but the cell population producing Gas6 in the bone marrow was never described. Likewise we find Adipor2/Adipoq(Adiponectin) among our interaction candidates with MSCs as main provider of adiponectin. Adiponectin stimulates HSCs proliferation and it was suggested that adipocytes could be its main source in the bone marrow [Masamoto et al., 2017]. Our analysis suggests that MSCs could be an important source of this cytokine for HSPCs.

Coexpression of ligand and antagonist by niche cells

Finally, our findings point to a role for ligand antagonists in the regulation of HSPCs. MSCs express high levels of both Cxcl12 and its antagonist Cxcl14 [Hara and Tanegashima, 2014]. Moreover, together with the expression of Bmp4, Bmp5, and Bmp6, MSCs also expressed Follistatin-1 (Fstl1), a transforming growth factor beta (TGFb) superfamily antagonist known for its inhibition of Bmp ligands. In the same manner, Endothelial cells express Jag1, a Notch1 and Notch2 ligand and Egfl7 which has been described as a Notch2 antagonist [Schmidt et al., 2009].

To summarize, using our interaction database and interaction scores, we could retrieve well known MSCs-HSCs interactions. We observed that niche/HSPCs interactomes were ordered firstly according to the niche component of the interactome, indicating that differences in ligands expressed by niche cells are higher than difference in receptors expressed by LT-HSCs, ST-HSCs and MPPs. Interactomes involving niche cells sorted in the endosteum (bone) clustered together, suggesting that cells from the endosteum tend to produce similar sets of ligands. Our method predicts that Gas6 and Adiponectin, two known regulators of HSPCS are specifically produced by MSCS. Finally our work suggests that ligands and their antagonists are sometimes produced by the same niche populations.

3.2 Quantification of cell cycle dynamics in vivo

Having worked on the identification of the molecular actors involved in niche-HSPCs interaction, we then focused on the characterization of cell cycle dynamics, a feature which can regulated by the niche in vivo. As mentioned in section 1.2, the quantification of cell cycle dynamics in vivo still remains an unmet goal.

We developed a method combining an EdU pulse-chase experiment and mathematical modeling to determine the length of each cell cycle phase as well as the cycling fraction within a cell population. Our experiment is based on work of [Baron and Penit, 1990], EdU is applied for a short period and the cells labeled during the pulse are then chased over a time period covering one to two cell cycle times so that cells first labeled in S phase transit to G2/M phase, then to G0G1 after completing their first cell cycle and then return to S phase as they enter a second cell cycle. We measure the position of EdU labeled cells in the cycle by flow cytometry and we use Bayesian inference to infer rate parameters from the experimental data.

3.2.1 Application of the method to a Cancer cell line: Tet21N

In order to test the validity of the method, we applied it to Tet21N human neuroblastoma, a cell line for which individual cell cycle times had been measured independently by time lapse microscopy by Erika Kuchen a former researcher in the lab (personal communication). We selected these cell based on their proliferative properties. They overexpress oncogene MYCN which forces them into proliferation [Ryl et al., 2017] and consequently we expect these cells to have no or very low G0 fraction.

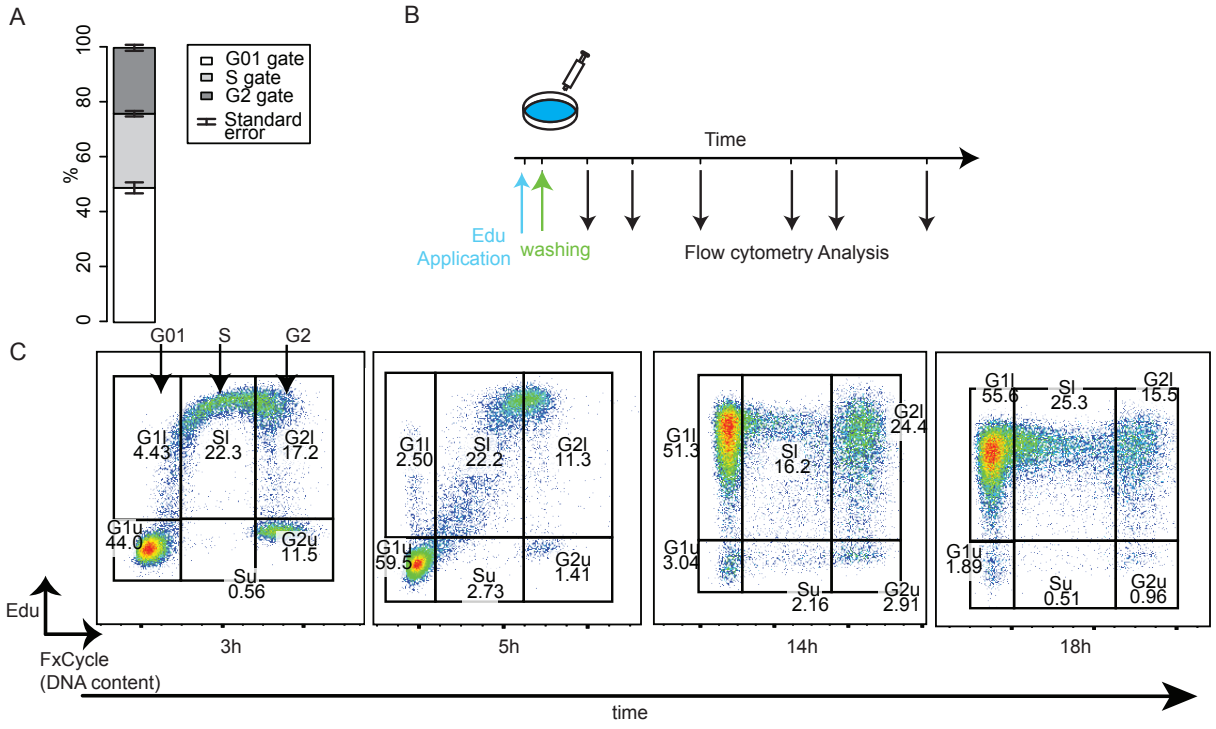


Figure 6: Analysis of Tet21N proliferation dynamics(A) Cell cycle distribution of Tet21N (B) Experimental protocol (C) Cell cycle progression of Tet21N during the pulse-chase experiment.

Cells in culture are treated with EdU and the plate is washed one hour after treatment (Fig. 6 B) so as to ensure that the experiment comprises both a pulse and a chase period. At multiple time points, cells are collected, the EdU content per cell is revealed by klick-it reation and the cells are stained for total DNA content. As shown figure 6 C, cell populations are divided according to total DNA in a standard way. We define a GO1 gate comprising cells with low DNA content, an S gate for cells with intermediate DNA content, and a G2M gate for cells with high DNA content. These gates are further divided in EdU positive ("l" for labeled) and negative("u" for unlabeled). 3h after injection, cells are labeled as they enter S phase. All cells in the S gate are labeled, and we also observe labeled cells in the G1 and G2M gates. It is well known that total DNA content staining cannot resolve perfectly cell cycle phases [Lacombe et al., 1988] and these labeled cells in G1 and G2M gate are likely to comprise cells in the S phase which can not be resolved as such cells based on full DNA staining. 5h After EdU application, we observe that new cells entering S phase are not EdU labeled (cells at the transition between the G1 and S phases gates are not labeled). This is an indication that EdU labeling efficiency has decreased. Over time, cells which have completed their

cell cycle return to G1, and as a result the fraction of labeled cells increases in the G01 gate. At the last time point (18h), we observe that the fraction of labeled cells has dropped in the G01 gate but increased in the S gate as a consequence of the return to the cycle of EdU labeled cells.

The scheme in figure 7 A represents the fictional journey of a single cell through the gates as it progresses through the cell cycle within the time scale of the experiment (in fact, each sample is an independent snapshot and we can only reconstitute trajectories by assuming that every samples will behave in the same manner over time). A cell, initially labeled in the S phase will be found in the G2M gate at a later time and upon division, its 2 daughter cells will be found in the G01 gate as the DNA content per cell is halved upon division. A recent study has shown that cell decision to enter a new cell cycle versus quiescence occurs at the end of G2M ([Spencer et al., 2013]).The progeny of our initially labeled cells can either reenter S phase after completing G1 or remain in the G01 gate beyond the length of duration of the G1 phase if it has become quiescent. By comparing the rate of entry into S phase of the total population to the rate of reentry into S phase of cells which were in S during the pulse we hope to estimate whether the general population contains a fraction which does not proliferate.

By measuring the progression of EdU labeled cells into G2M, then to G1 and finally to S phase again, we can estimate the full cell cycle length. At steady state, the cell cycle distribution between the G01 gate, the S gate, and the G2M gate depends on the length of each phase, by estimating the length of G1, we can predict whether the cells in the G01 gate are all proliferating or whether a fraction of the cells within this gate does not actually cycle.

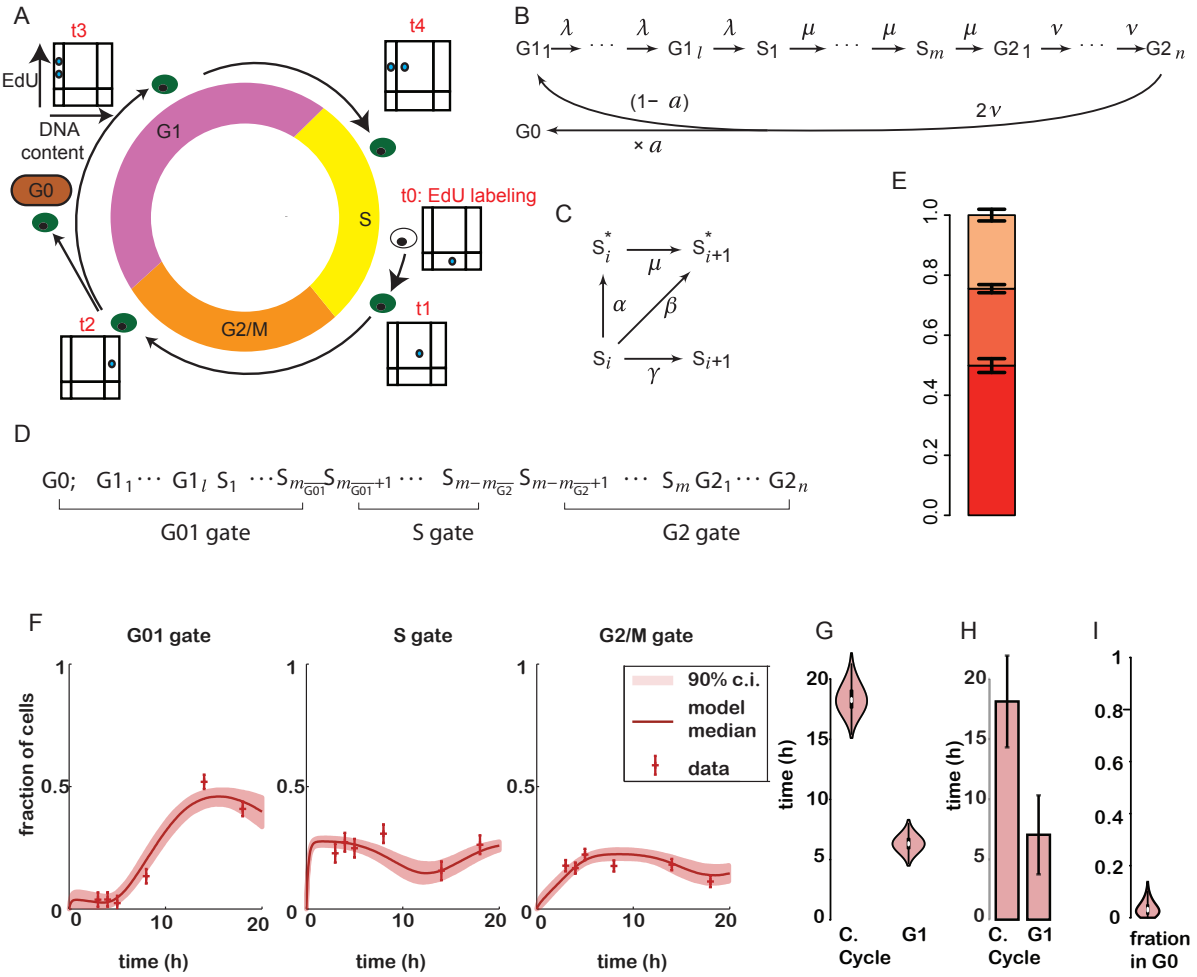


Figure 7: Modeling cell cycle dynamics from EdU pulse-chase experiment (A) cell cycle scheme (B) mathematical model of cell cycle progression (C) model of EdU incorporation in S phase (D) assignements of the G01, S and G2M gates to the model (E) model prediction of steady state fraction of cells in the G01, S and G2M gates, error bars represent 90% credible interval from the posterior distribution (F) time course of EdU labeled cells in G01, S and G2M gates in Tet21N, data are mean values +/- standard errors (G) cell cycle length and G1 length estimated from model (H) cell cycle length and G1 length of Tet21N measured by time lapse microscopy (I)fraction of cells in G0 in Tet21N as estimated by the model.

3.2.2 Mathematical model of the cell cycle

Based on these elements, we formulated a mathematical model to represent cell cycle progression and EdU incorporation. This mathematical model was developed with the help of Nils Becker, a senior postdoc in the Höfer Lab.

The model is presented in figure 7 B. We assume that, with respect to the cell cycle, a cell can be in 4 states : G0 (cell cycle arrest), and the 3 cell cycle phases G1, S and G2M. Upon division, a fraction a of the divided cells will enter quiescence (G0) while the rest reenter G1.

The variance in length of each cell cycle phase is expected to be lower than that of an exponential distribution which would apply if we chose to model each cycle phase as a single state in a linear population model. For an exponential distribution, the standard deviation is equal to the mean and the length of a specific phase can be infinitely small. For the S phase, this would imply that DNA replication can be completed instantly by a cell upon entering the phase. In order to account for the existence of a minimum length for each cell cycle phase We divide the G1, S and G2 phases in l, m and n substeps respectively (Fig. 7 B). The progression through the substeps approximates the progressions of the cells through the phases.

The progression of cells between two substeps is exponentially distributed. The distribution of a sum of k exponential variable β is an Erlang distribution. The coefficient of variation of this distribution is given by $1/\sqrt{k}$ and the mean is given by $\beta/\sqrt{k}$. The choice of the Erlang distribution to model the variability of cell cycle length was motivated by computational convenience but is also supported by recent findings [Chao et al., 2019].

Cells progress within the G1 phase with a transition rate λ so that the mean length of G1 is given by l/λ The rate of progression in S is μ and the mean length of S is m/μ The rate of progression in G2M is ν with a mean duration of n/ν. Upon completion of G2M, a fraction a of the dividing cells will enter G0 while a fraction $1-a$ will enter G1 and continue to cycle.

We consider that cells can be lost from any cell cycle phase with a rate δ. This rate reflects both loss by death and differentiation.

the system can be described by the following differential equations:

$$
\begin{aligned}
&\tfrac{d}{dt}G1_1(t) = 2\,G2_n(t)\nu(1-a) - G1_1(t)(\lambda+\delta), \\
&\tfrac{d}{dt}G1_i(t) = G1_{i-1}(t)\lambda - G1_i(t)(\lambda+\delta); && 2 \le i \le l \\
&\tfrac{d}{dt}S_1(t) = G1_l(t)\lambda - S_1(t)(\delta+\mu), \\
&\tfrac{d}{dt}S_i(t) = S_{i-1}(t)\mu - S_i(t)(\delta+\mu), && 2 \le i \le m \\
&\tfrac{d}{dt}G2_1(t) = S_m(t)\mu - G2_1(t)(\delta+\nu), \\
&\tfrac{d}{dt}G2_i(t) = G2_{i-1}(t)\nu - G2_i(t)(\delta+\nu), && 2 \le i \le n \\
&\tfrac{d}{dt}G0(t) = 2\,G2_m(t)\nu a - G0(t)\delta.
\end{aligned}
\tag{3.7}
$$

Steady Growth

We now consider a population of cell in culture which is growing exponentially with a rate κ, so that $\frac{d}{dt}C = C\kappa$. In such a population in steady growth, the number of cell in each cell cycle phase has to grow exponentially with the same rate κ so that the fraction of cells in each compartment remain constant. In this case, we can write the equality below:

$$
\begin{aligned}
&\tfrac{d}{dt}G1_1(t) = 2\,G2_n(t)\nu(1-a) - G1_1(t)(\lambda+\delta) = G1_1(t)\kappa, \\
&\tfrac{d}{dt}G1_i(t) = G1_{i-1}(t)\lambda - G1_i(t)(\lambda+\delta) = G1_i(t)\kappa; && 2 \le i \le l \\
&\tfrac{d}{dt}S_1(t) = G1_l(t)\lambda - S_1(t)(\delta+\mu) = S_1(t)\kappa, \\
&\tfrac{d}{dt}S_i(t) = S_{i-1}(t)\mu - S_i(t)(\delta+\mu) = S_i(t)\kappa, && 2 \le i \le m \\
&\tfrac{d}{dt}G2_1(t) = S_m(t)\mu - G2_1(t)(\delta+\nu) = G2_1(t)\kappa, \\
&\tfrac{d}{dt}G2_i(t) = G2_{i-1}(t)\nu - G2_i(t)(\delta+\nu) = G2_i(t)\kappa, && 2 \le i \le n \\
&\tfrac{d}{dt}G0(t) = 2\,G2_m(t)\nu a - G0(t)\delta = G0(t)\kappa.
\end{aligned}
\tag{3.8}
$$

This system can be rewritten in matrix form :

$\frac{d}{dt}C = \mathbf{T}C = C\kappa$, where (3.9)

$C = [G1_1(t), \ldots, G1_l(t), S_1(t), \ldots, S_{1m}(t), G2_1(t), \ldots, G2_n(t), G0(t)]^{\mathsf{T}}$, and

$$\mathbf{T} = \begin{bmatrix}
-(\lambda+\delta) & & & & & & & & & & 2(1-a)\nu & 0 \\
\lambda & \ddots & & & & & & & & & & \vdots \\
 & \ddots & -(\lambda\delta) & & & & & & & & & \\
 & & \lambda & -(\mu+\delta) & & & & & & & & \\
 & & & \mu & \ddots & & & & & & & \\
 & & & & \ddots & -(\mu+\delta) & & & & & & \\
 & & & & & \mu & -(\nu+\delta) & & & & & \\
 & & & & & & \nu & & & & & \\
 & & & & & & & \ddots & \ddots & & & \\
 & & & & & & & & \nu & -(\nu+\delta) & \vdots & \\
 & & & & & & & & & 2a\nu & -\delta &
\end{bmatrix}. \tag{3.10}$$

κ is an eigenvalue of the matrix T with the state vector C its associated eigenvector. It must be the eigenvalue with the highest real part since only that one survives at steady growth.

We observed that this system accepts at most only one positive real eigenvalue. This eigenvector can then be normalised to one in order to compare the model values to the experimental data which are given as fractions.

Steady State

Our method is designed to be applied to cell cycle in vivo. In the adult mouse, in an homeostatic context, cell population can be considered at steady state, so that the total cell number does not change in time. In this context, the growth rate $\kappa = 0$.

We define $\overline{G0}$,$\overline{G1}$,$\overline{S}$,$\overline{G2M}$, the steady states of the phases G0, G1, S and G2M respectively. The cells produced by cell division at the end of the G2M phase must be equally compensated by the death rate δ So that $\overline{G2M}_n \nu = C\delta$

If we normalize the population to $C = 1$, then $\overline{G2M}_n \nu = \delta$

We can then find the steady state values of each cell state by back-substitution:

$$\overline{G2}_1 = \overline{G2}_n \frac{(\nu+\delta)^{n-1}}{\nu} \tag{3.11}$$

$$\overline{G2} = \sum_{i=1}^{n-1} G2_n \frac{(\nu+\delta)^i}{\nu} = \overline{G2}_n \frac{1-(\frac{\nu+\delta}{\nu})^n}{1-\frac{\nu+\delta}{\nu}} \tag{3.12}$$

$$\overline{S}_m = \overline{G2}_1 \frac{(\nu+\delta)}{\mu} \tag{3.13}$$

$$\overline{S}_1 = \overline{S}_m \frac{(\mu+\delta)^{m-1}}{\mu} \tag{3.14}$$

$$\overline{S} = \sum_{i=1}^{m-1} \overline{S}_m \frac{(\mu+\delta)^i}{\mu} = \tag{3.15}$$

$$\overline{S}_m \frac{1-(\frac{\mu+\delta}{\mu})^m}{1-\frac{\mu+\delta}{\mu}} \tag{3.16}$$

$$\overline{G1}_l = \overline{S}_1 \frac{(\mu+\delta)}{\lambda} \tag{3.17}$$

$$\overline{G1} = \sum_{i=1}^{l-1} \overline{G1}_l \frac{(\lambda+\delta)^i}{\lambda} = \overline{G1}_l \frac{1-(\frac{\lambda+\delta}{\lambda})^l}{1-\frac{\lambda+\delta}{\lambda}} \tag{3.18}$$

$$\overline{G0} = \overline{C} - (\overline{G1} + \overline{S} + \overline{G2}) \tag{3.19}$$

$$a = \overline{G0}/2; \tag{3.20}$$

EdU incorporation

We can observe that EdU incorporation is such that cells in S phase will initially become EdU labeled with a very high efficiency, so that after one hour all cells in S phase are labeled but the EdU concentration in the medium will decrease over time until new cell entering S phase never get above the label detection threshold. A cell needs to replicate a certain quantity of DNA in order to become EdU positive. However, the precise amout of EdU that needs to be incorporated so that a cell becomes EdU positive is unknown and the time spent in S phase that is necessary to become EdU positive is also unknown. Consequently, we choose to model EdU incorporation in cycling

cells as a process α, independent and happening in parallel with cell cycle progression. During the S phase a cell can become EdU positive within a substep with a rate ϕ or as it progresses to the next substep with a rate ν, so that $\alpha = \phi + \nu$ (Fig. 7 C).

As mentioned previously, cells transit from one substep to the next with a rate β. The EdU incorporation does not affect the rate of progression of cells from one compartment to the next, so that $\beta = \phi + \zeta$, where ζ is the progression of the non labeled cells from one substep to the next.

The EdU concentration has been described before has decaying exponentially in mice [Cheraghali et al., 1994], our main model of interest. Therefore, we model EdU incoporation as an exponential function $\alpha = \alpha_0 \exp(-t/\epsilon)$. Where ϵ is a free parameter that we estimate and α_0 is fixed.

3.2.3 Acurate prediction the cell cycle duration of Tet21N

As shown in Fig. 7 C), our model can be well fitted to the observed steady state distribution of cells in the cycle presented in Fig. 6 A. The fitted kinetics of EdU labeled cells in each of the gates over time also agrees with the data (Fig. 7 D), we observe that initially, cells are only labeled in S phase and that the labeled phases then progresses to the other cell cycle phases while the fraction of labeled cells decreases in S following EdU dilution. The oscillatory behaviour of these curves depends on the variability in the length of each cell cycle phases. The more determined is the length of each phase, the more it imposes delays in the progression of the EdU label.

The mean cell cycle length in the range of 16.7 to 20.3 hours (90% credible interval) 7 is consistent with direct observations from time-lapse microscopy on the FUCCI marker expressing Tet21N (Fig.7 G) and so is the lengh of the G1 phase. This results validate the method.

Furthermore, our model predicts that the fraction of quiescent cells in the Tet21N is comprised between 0 and 5% (Fig.7 I, Sup. Fig.6), as we expected from these cells overexpressing MyCN.

3.2.4 Quantification of cell cycle dynamics of Double Positive Thymocytes

We have shown in the previous part that our method can quantify cell cycle dynamics in vitro, here we apply it to a cell population in vivo, the murine Double Positive (DP) thymocytes.

DP have been described as containing both a quiescent and a cycling population although there is no consensus on the actual fraction of cycling cells [Mingueneau et al., 2013]. These cells undergo selection based on their T cell receptor affinity for MHC+peptide resulting in 90% of cell death by neglect. They receive input from less differentiated thymocytes, but they make up the large majority of the thymocytes (80% to 90%) while their direct upstream progenitors, the Immature Single positive (ISP) make up 0.26% of the mouse thymocytes ([Gegonne et al., 2018]). Therefore, the differentiation rate of ISP should be a 100 fold higher than the proliferation rate of DPs to affect significantly the EdU label dynamics in DPs. As a consequence ISPs are very unlikely to have any significant impact over the EdU label dynamics of DPs within the time of our experiment (18h). For this reason, the DP offer a good system to apply our model which does not include influx. For the purpose of this work, DP were always FACS gated in a standard way as described in Sup.Fig 3.

BrdU has been previously described to be toxic and to affect cell cycle in vivo [Wilson et al., 2008]. Wilson and colleagues describe that upon prolonged treatment with BrdU in drinking water, the cell cycle distribution was affected, with a higher fraction in the S and G2M phases which they interpreted as the increased proliferation. In order to ascertain that cell cycle distribution was not affected by our treatment with EdU, we measured cell cycle distribution of non treated mice in 3 separate experiments and then compared the cell cycle distribution to our EdU treated mice, for that comparison, we pooled EdU data from EdU treated mice, reasoning that such a perturbation would change cell cycle distribution during the whole time of our assay. We did not notice any significant difference between control and EdU treated mice (Sup.Fig 4); this confirmed that we could interpret our EdU data as being representative of unperturbed steady state cell proliferation.

It has been proposed that early DP are actively proliferating and then differentiate into quiescent late DP, and it is generally considered that the transition to quiescence occurs before positive

selection [Yates, 2014]. Based on cell cycle distribution, DP certainly appears lowly proliferating, with more than 90% of cells in the G01 gate (Fig. 8 A)

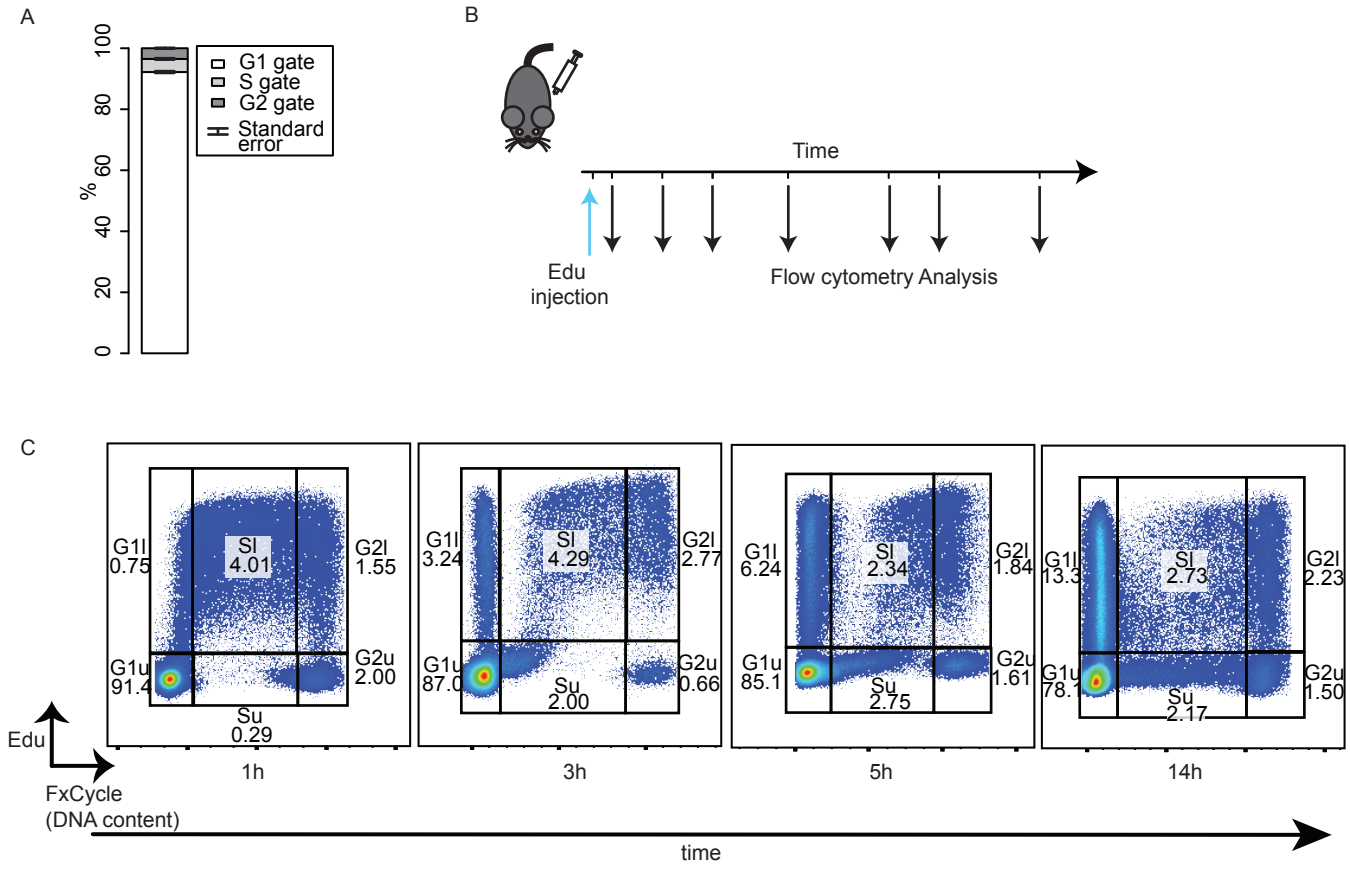


Figure 8: Analysis of Double Positive thymocytes proliferation dynamics(A) Cell cycle distribution of DP (B) Experimental protocol (C) Cell cycle progression of DP during the pulse-chase experiment

For this in vivo experiment, we inject EdU intra peritoneally and then obtain samples over the course of 18h (Fig. 8 B). As anticipated, EdU is degraded during the first hours following injection so that, 3h hours after injections, new cells entering the S phase are not labeled (Fig. 8 C). After 3 hours, we already observe that some cells which have been labeled in S phase initially have returned to G1, indicating that some cells can complete G2M in less than 3h. 5 hours after EdU injection we start to observe that cells which have returned to G1 enter the S phase gate, and 14h after injection, the S and G2M gates are now filled with EdU labeled cells.

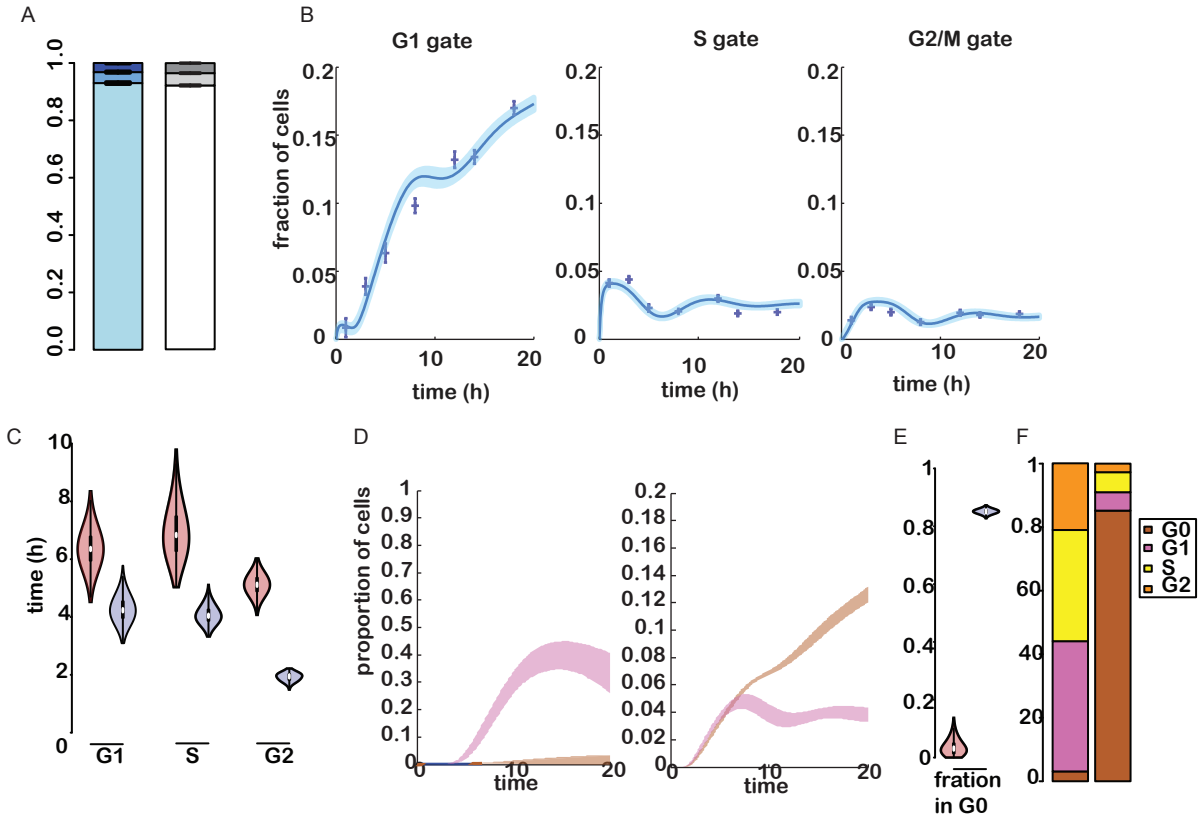


Figure 9: Cell cycle Model predictions (A) comparison between model predictions (left barplot) and measurement (right barplot) for steady state fractions of cells in G0G1 gate (lower layer), S gate (middle layer) and G2M gate (top layer), error bars are 90% credible interval for the model and standard error to the mean for the measurements(B) fraction of EdU label cells in each gate over time in DP, light blue area corresponds to the 90% credible interval (C) model predicted Cell cycle phase length for Tet21N (red) and DP(blue) (D) predicted kinetics of EdU labeled cells in G1 phase (pink) and G0 phase(brown) within the G01 gate in Tet21N (left) and DP(right) (E) steady state fraction of cells in G0 in Tet21N(red) and DP(blue) (F) median predicted cell cycle distribution in Tet21N(left) and DP(right)

The observed cell cycle distribution and EdU label kinetics are reproduced well by our model(Fig. 9 A, B).

At face value, the high fraction of cells in the G01 gate (Fig.9 A) could be interpreted as a very long G1 phase relative to S and G2M. Here, our novel method involving EdU labeling and mathematical modelling allows us to discover that the length of G1 in DP is actually very similar to that of the S phase (4h and 3.9h respectively) (Fig. 9C), and that most cells (85%) within the G01 gate are actually non cycling cells (Fig. 9E). By contrast, our analysis reveals that DPs, with a very low fraction in S and G2M compared to the Tet21N actually comprise a fast cycling population, with a

mean cell cycle time inferior to 10 hours, almost twice as fast as the Tet21N Fig. 8 C). In fact, the length of each of the three phases G1, S and G2M is higher in Tet21N. The model can separate EdU labeled cells in the G0 phase from those in G1, it predicts that a large fraction of the EdU labeled DP which accumulate in the G01 gate are in fact G0 cells (Fig. 9 D). While the fraction of labeled G1 cells in the G0G1 gate approaches a plateau after a few hours(just like EdU labeled cells in S and G2M gates), the fraction of G0 cells keep on increasing during the remainder of the experiment. This large accumulation of labeled cells in G0 is a consequence of the large fraction of cells (42%; as described in section 3.2.2, a is equal to 0.5 times the G0 steady state fraction) which transit to quiescent state upon division. By contrast Tet21N, very few labeled G0 cells (if any) accumulate in the G01 gate consistently with a parameter a comprised between 0 and 2%.

The large predicted quiescent fraction in DPs and the unidirectional transition from a cycling to a quiescent state are consistant the classical subdivision of DP into early DP (cycling) and late DP (quiescent) (see for instance [Mingueneau et al., 2013, Yates, 2014]).

Of note, the variability between samples, is very low for the thymocytes compared to the Tet21N as demonstrated by the standard errors to the mean (the standard errors are comprised between 0.03 and 0.05 for Tet21N versus 0.0025 to 0.006 for the DPS). Mammalian cells reside normally in a very controled *milieu interieur*, where temperature and gas concentrations vary very little. Such is not the case for cells growing in a culture room. A cell incubator used frequently by several experimentalists will have high fluctuation in temperature, C02 and 02 concentrations. One can not expect as little variation as between individual mice.

In summary, using a simple and fast experiment involving a standard biological technique, we can infer the length of each cell cycle phase and the quiescent fraction in vitro as well as in vivo via mathematical modeling and Bayesian inference. the cell cycle method recovers accurately the cell cycle length of a cancer cell line in vitro. It further predicts that the DPs contain a highly proliferative subpopulation and that close to 50% of these cycling cells transit to a state of quiescence upon division, a process which results in a majority of quiescent double positive.

3.3 Regulation of differentiation in early thymocytes: lessons learned from thymus autonomy

As described in introduction, thymus autonomy in absence of progenitors, as observed by V. Martins and colleagues [Martins et al., 2012, Martins et al., 2014], was explained by competition for limited resources between progenitors (Il-7), so that in absence of competitors from the bone marrow the local progenitor were able to remain in the compartment. In their 2014 paper Martins and colleagues had further suggested that bone marrow progenitors had a proliferative advantage over the resident thymocytes to explain why they constantly replace resident thymocytes under physiological conditions.

Thymus autonomy can only be achieved in a Il2rγ (or Il7ra), Rag2 K0 mouse. Il2rγis notably a component of the Il-7 receptor a critical cytokine involved in proliferation and differentiation at the DN2 and DN3 stages of T cell development. Furthermore, the recombination of the beta chain occurs at the DN2-DN3 stage and cells lacking a recombined beta chain cannot progress beyond this stage. These two facts taken together made the DN3 stage an early candidate "thymic stem cell" in the autonomous thymus [Boehm, 2012].

In fact, phenotypical observations also suggest that DN3 is the likely compartment where the competition occurs and therefore a candidate "thymic stem cell" under thymus autonomy . Indeed, while most major T cell progenitors populations are represented in the autonomous thymus (Fig. 10), they seem to lack DN2 progenitors (Fig. 10).

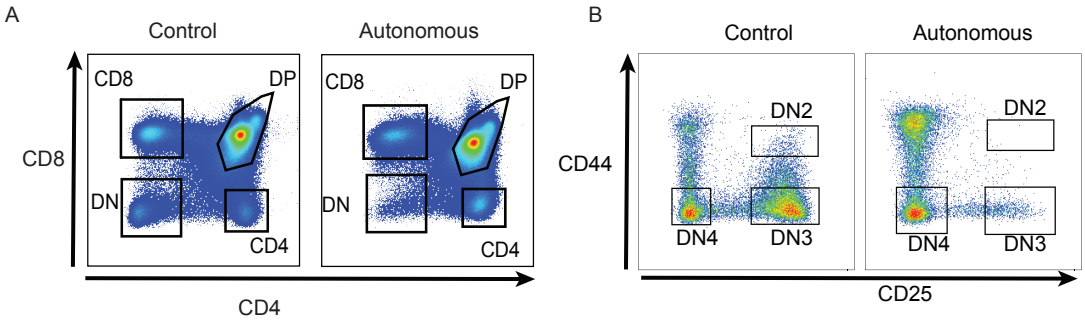


Figure 10: Thymocytes populations in control and autonomouy thymi (A) total thymocytes (B)double negative thymocytes

3.3.1 Mathematical model of thymic competition

In order to understand this competition process, we decided to model it and simulate the conditions of the experiment. Population growth limited by ressources is classically represented in ecology by a Verhulst logistic model [DeBoer, 2020]. If we assume a population N growing with a proliferation rate r, where the growth of the population is restricted by the then the proliferation can be written as :

$$\frac{dN}{dt} = Nr(1 - \frac{N}{K}) \tag{3.21}$$

In such a system, a population grows with an ever decreasing growth rate until its size reaches the carrying capacity of the system, at which point the population ceases to grow.

Now, we consider a population of cells, the "DN3" receiving an influx from the upstream cell compartment, it proliferates with a rate λ and differentiates with a rate α.

We can consider that proliferation is constrained by resources in the same manner as in equation 3.21

so that:

$$\frac{dDN3}{dt} = p - DN3\lambda(1 - \frac{DN3}{K}) - DN3\alpha \tag{3.22}$$

Likewise, the differentiation could be dependent on the capacity of the system so that it increases as the size of the system gets close to saturation:

$$\frac{dDN3}{dt} = p - DN3\lambda - DN3\alpha\frac{DN3}{K} \tag{3.23}$$

In both cases, the size of the population is controled by the capacity of the system, either by reducing proliferation or increasing differentiation.

For a number of cells at steady state DN3=100000 and the following parameters $\alpha = 0.25/day$ an $\lambda = 0.55/day$, we compute the values for p at steady state, respectively $\alpha DN3 - \lambda DN3(1 - \frac{DN3}{k})$

for eq. 3.22 and $\alpha DN3(+\frac{DN3}{k}) - \lambda DN3$ for eq. 3.23.

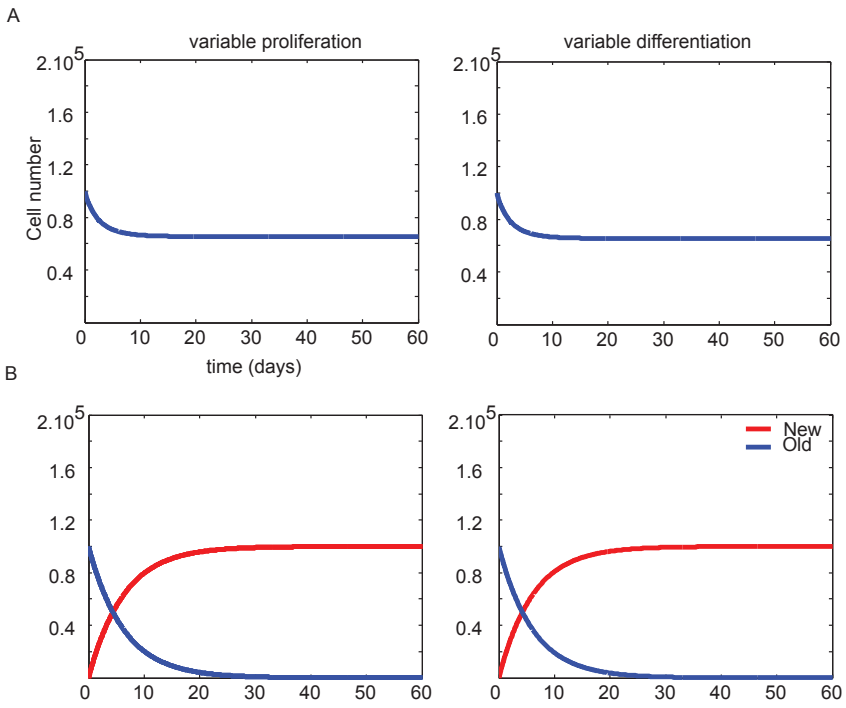


Figure 11: Simulation of thymocyte dynamics according to the competition models left panels proliferation dependent on competition. right panels differentiation dependent on competition A. no competition (autonomous thymus) B. competition (control)

When simulating the behaviour of the two systems with identical sets of parameters, and after setting the p=0 (absence of influx from upstream progenitors) (Fig: 11 A). One can observe that the two models behave in the same fashion. While the two curves do not really overlap, we see a drop from steady state and then a plateau as a new steady state dependent purely on α and λ is established. The two models can account for the behaviour of the DN2/DN3 cell population in thymus autonomy.

Using these two models, we can also simulate the influx of bone marrow progenitors in the thymus graft upon transplantation in a wild type mouse.

If one now consider that old and new cells share the same resource k, and that only the New cells benefit from influx . For the case where λ depends on competition, one can write :

$$\frac{dDN3_{Old}}{dt} = DN3_{Old}\lambda(1 - \frac{DN3_{Old} + DN3_{New}}{K}) - DN3_{Old}\alpha$$
$$\frac{dDN3_{New}}{dt} = p - DN3_{New}\lambda(1 - \frac{DN3_{Old} + DN3_{New}}{K}) - DN3_{New}\alpha$$

And in the case where differentiation is increased by competition:

$$\frac{dDN3_{Old}}{dt} = DN3_{Old}\lambda - DN3_{Old}\alpha(\frac{DN3_{Old} + DN3_{New}}{K})$$
$$\frac{dDN3_{New}}{dt} = p - DN3_{New}\lambda - DN3_{New}\alpha(\frac{DN3_{Old} + DN3_{New}}{K})$$

Figure 11 B shows that both models can indeed qualitatively reproduce the replacement occuring in the wild type thymus. Interestingly thes models do not assume any advantage of the bone marrow progenitor versus resident thymocytes, argument that was raised by V.Martins and colleagues to explain the the systematic replacement of resident thymocytes. Mathematical modeling reveals that influx is sufficient to ensure constant replacement of resident population.

It follows from these simulations that cell maintenance in absence of competition can either result of increased proliferation or reduced differentiation.

At the DN3 stage occurs the first major checkpoint in development of at T cell. At this stage, the β chain of the T cell receptor(TCR) is rearranged. This process happends in two steps, first a Joining (J) segment is randomy joined to a Diversity (D) segment on the DNA locus while the genomic region in between these segments is removed. Then, a V segment is joined to the newly formed DJ region. At each of these two steps, nucleotides are introduced and removed randomly at the junction cites between these V, D and J segments. These randomly introduced nucleotides increase the diversity of the chains at the cost of creating two sequences out of frame for every three recombi-

nations. These out of frame sequences cannot produce functional proteins. This process occurs in parallel on both the alleles; upon productive recombination, the β chain is expressed and forms the pre-TCR with a constant pre alpha chain (Ptrca) at the surface of the cell. The activation of the pre-TCR leads to further differentiation into the T cell lineage [Michie and Zuñiga-Pflücker, 2002].

Based on β chain intracellular expression, DN3 can be subdivided into DN3a and DN3b. DN3a complete the TCR β recombination, and DN3b are cells which have been selected via the pre-TCR signal. DN3a are maintained in a quiescent state as the cells complete the β chain recombination and the chains are tested for productive recombination [Boudil et al., 2015].

Upon activation of the pre-TCR and costimulation by Notch ligands and Il-7 [Boudil et al., 2015], DN3a transit to the highly proliferative DN3b. It has been shown that these signals induce both proliferation and differentiation in a coupled manner, so that proliferation is thought to be necessary and sufficient to induce differentiation [Kreslavsky et al., 2012].

This strong coupling of the processes implies that increased proliferation rate at this stage should not rescue a thymus lacking bone marrow progenitors. If DN3 are indeed the new "thymic stem cells" under autonomy, this can happen through two distinct phenomena: the proliferation of the normally quiescent DN3a population, or the decoupling of proliferation and differentiation at the DN3b stage.

3.3.2 Phenotypic DN3 are proliferating in Thymus autonomy

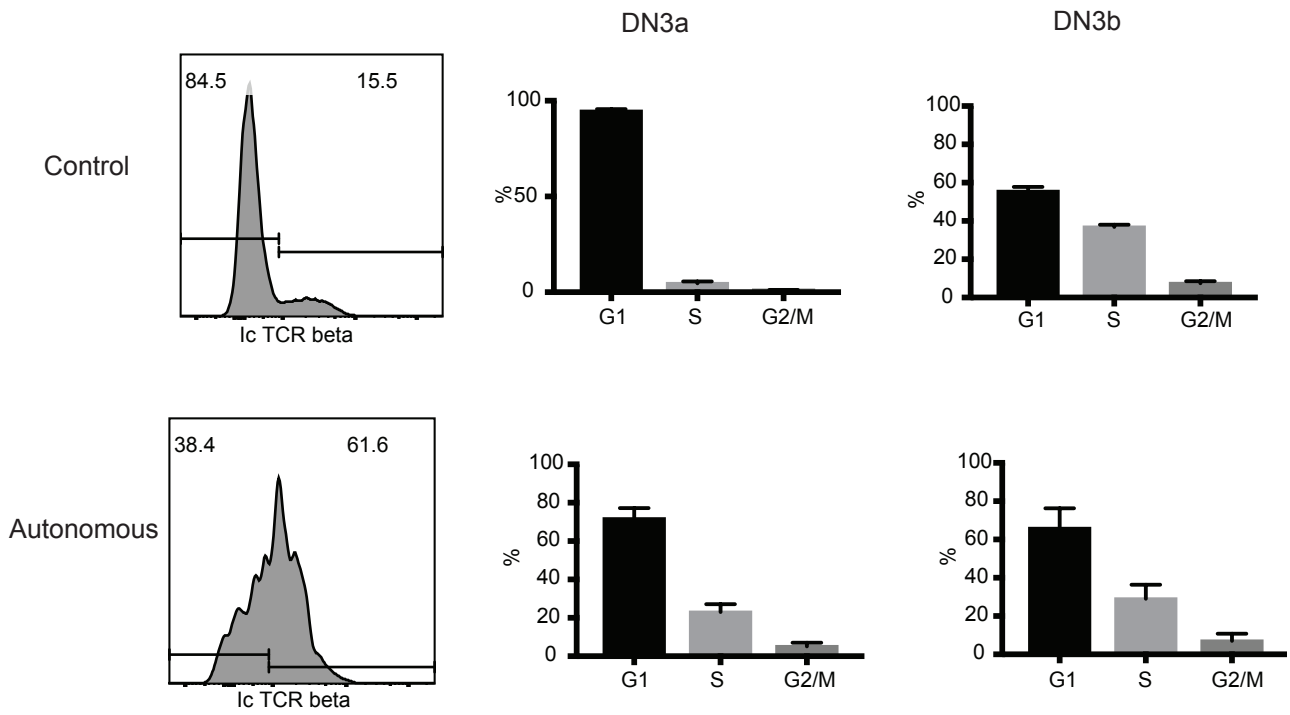


Figure 12: Gating of DN3a and DN3b thymocytes based on intra cellular Beta chain expression and mean distribution of cells in cell cycle based on total DNA staining in both populations in control and automous thymi (4 weeks after transplantation)

Figure 12 shows that in the control thymus, the intracellular β chain expression allows a clear segregation of two populations. We note that in the DN3a, which do not express the chain, very few cells are in S and G2M gates, this contrasts whith the β chain expressing DN3b which have in average more than 30% of cells in S and G2M gates. This result is consistent with the idea that DN3a are mostly quiescent and that TCR β expression is associated with higher proliferation of these cells. In the autonomous thymi however we cannot separate two distinct populations based on TCR beta expression. Nevertheless, when gating TCR beta low versus TCR beta high, one can observe that TCR β low cells have a lower fraction of cells in S an G2M than the TCR β high, suggesting that TCR β low and high mighst still have distinct proliferation characteristics. One can conclude, that consistantly with what was observed by [Peaudecerf et al., 2012], these phenotypic DN3 are proliferating in the autonomous thymus which makes them a serious thymic stem cells contender in the autonomous context but it still uncertain at this point whether the DN3 in the autonomous thymus comprise only DN3b or whether it still contains cell without recombined β

chain.

3.3.3 Single cell transcriptome reveals a perturbation of differentiation at DN3 stage

We used single cell RNA seq to describe the DN3 population on a molecular level. We sorted DN3 from endogenous and autonomous thymi 4 weeks after transplantation, and performed single cell RNAseq using 10X Genomics technology with targeted enrichment of T cell receptor sequences. The transcriptome was aligned to the mouse genome and the T cell receptor VDJ sequences were reconstituted and annotated with Cell Ranger, software of 10X Genomics. We obtained 800 to 1200 good quality cells per sample after quality control performed according to the Bioconductor workflow [Lun et al., 2016].

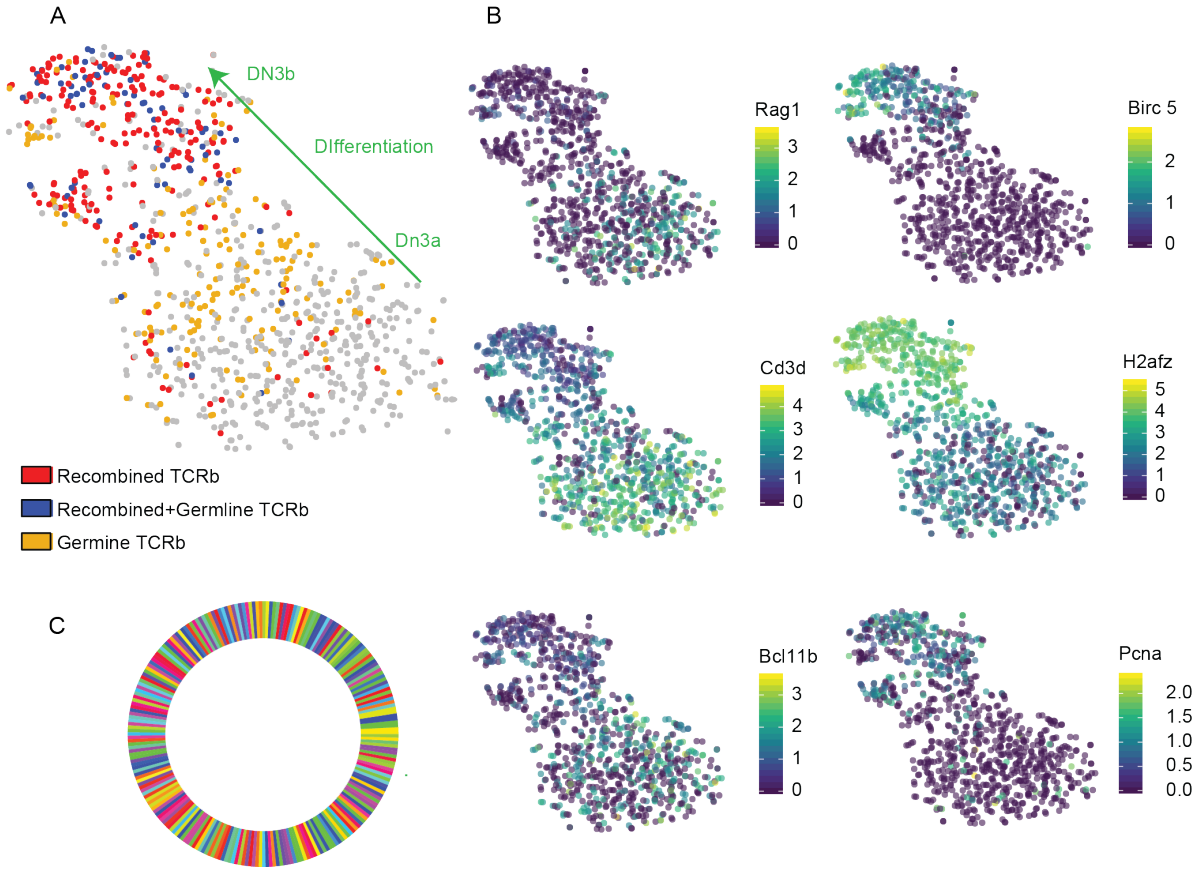


Figure 13: Single Cell transcriptome and TCR beta chain repertoire diversity of DN3 thymocytes in control .(A) t-SNE plot of single cell transcriptomes from input-dependent(DN3 (B) Log2 expression values of genes involved in T-cell differentiation at the DN3 stage (C) Clone sizes of recombined β chains in both conditions

As expected from DN3 progenitors, almost no cells (0 to 1 cell per sample) were found to express a recombined α chain, which is recombined after the DN3 stage.

We found that 26% of cells expressed rearranged β chains (CDR3 sequence detected). On the TSNE plot (Fig.13 A), we observe that these cells cluster together based on their transcriptome.

For many cells β chain transcripts were detected but no CDR3 sequence. It is established that TCR recombination requires the transcription of germline segments as the transcription favors accessibility of the locus to the recombinases [Sleckman et al., 1996]. We checked whether some of the sequences without CDR3 were actually germline transcripts.

The region comprised between the J segments and the D segment of the TCR loci are systematically removed during the recombination process when D and J are joined. Using BLAST, we checked whether some of the TCR transcripts without CDR3 contained such germline sequences. We found a significant fraction of cells expressing germline transcripts Cells expressing exclusively germline transcripts (orange cells) are well separated from the cells expressing rearranged β chains (blue and red). The full transcriptome thereby confirms that these cells form a distinct population and supports the use of germline TCR expression as a marker for ongoing rearrangement.

In 13 B, we present the expression of a few genes known to be involved in early T cell development. We observe the pattern of expression of Rag1, a recombinase involved in TCR rearrangement is higher in cells which do not express a recombined β chain, likewise CD3e, a component of the TCR signaling machinery is overexpressed in these cells. These cells overexpress Bcl11b, a transcription factor which is known as an hallmark of T cell commitment at the DN2 and early DN3 stages ([Liu et al., 2010]) (early DN1 and DN2 still have NK cell potential). Taken together, the expression pattern of these genes validate these cells as early DN3.

The cells carrying recombined β chains overexpressed Birc5, a gene coding for an antiapoptotic protein expressed by cells having gone through β selection [Xing et al., 2004]. They also overexpressed H2afz, an histone subunit, is an hallmark of proliferation in thymocytes and Pcna [Puthier et al., 2004], a classical marker of S phase. This expression pattern point to a more proliferative population, as the β selected cells should be. We can conclude that single-cell transcriptomic provides a fine-grained picture of β selection.

TCR targeted sequencing gives us a glimpse of the β chain repertoire (Fig. 13 C). The sampled repertoire shows extreme diversity with only two clones being sampled more than once (twice). This is consistent with a short residence time of cells carrying a recombined β chain in the DN3 compartment as enforced by the preTCR signaling. although the sampling effect might hide the truth.

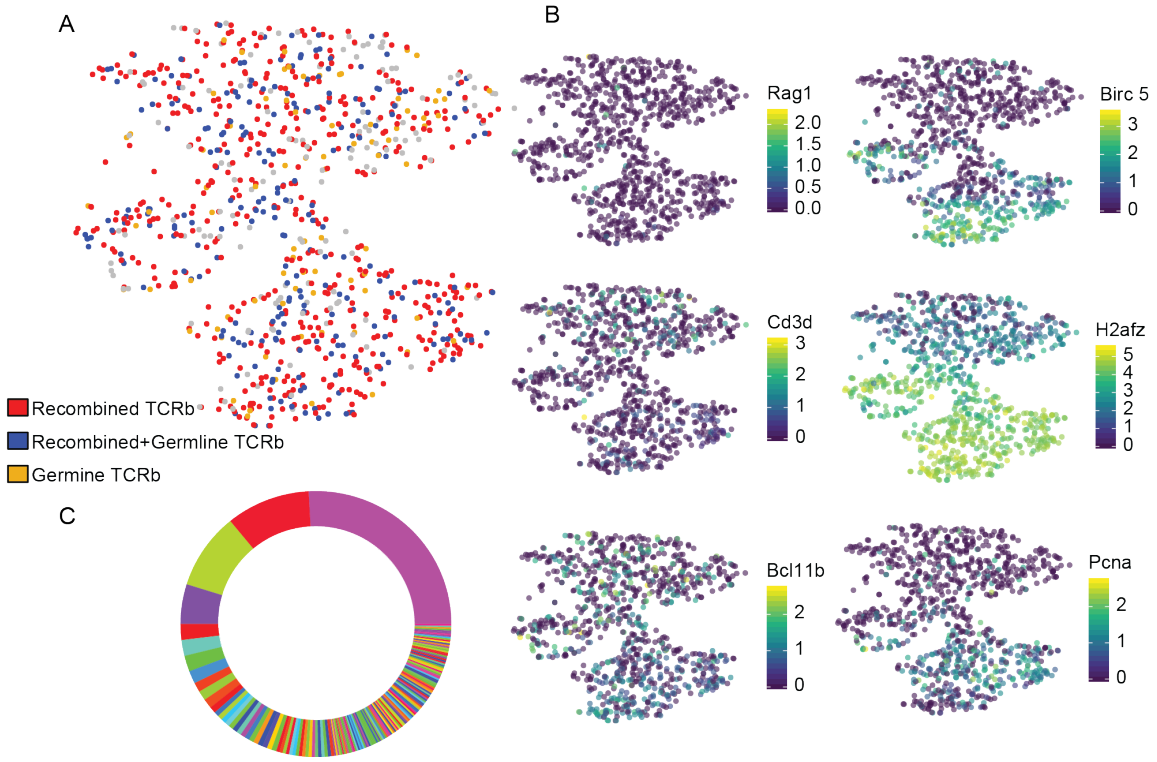


Figure 14: Single-cell transcriptome and TCR beta chain repertoire diversity of DN3 thymocytes in autonomous thymi .(A) t-SNE plot of single cell transcriptomes from input-dependent(DN3 (B) Log2 expression values of genes involved in T-cell differentiation at the DN3 stage (C) Clone sizes of recombined β chains in both conditions

By contrast, we observe that this pattern of differentiation is lost in the autonomous thymus (Fig. 14 A). There seem to be no association between TCR status and transcriptomic state anymore. We can still note that, while the majority of the cells express a recombined β, some cell still express exclusively a germline TCR. This could indicate that recombination is still going on to some extent in the autonomous thymus. The expression pattern of the six genes already introduced in figure 13 shows that the cells in the lower part of the TSNE plot express Pcna and H2afz, genes associated with proliferation. Here we find no association between proliferation and TCRb recombination status which confirms the disconnection between β selection and proliferation.

Like in the control, Figure 14 C shows the TCR β repertoire for the cells where a productive β

chain was identified. Each color is a distinct TCR clone and the thickness of each stripe reflects the relative size of the clone. The diversity of the beta chain repertoire in the input-dependent DN3 is representative of a population with strict coupling of proliferation and differentiation. We observed very little amplification of the TCR clones, almost every TCR β chain is uniquely represented. The repertoire of the autonomous thymus reveals a very different situation. While a large TCR diversity persists, a few clones are massively expanded, one clone representing 15% of the overall repertoire. This indicates that cells with a productive β chain now expand without leaving the DN3 compartment. We can conclude that the coupling between differentiation and proliferation is abolished in the autonomous thymus. This observation challenges the common understanding that beta selection induces necessarily the progression in the αβ T cell lineage.

3.3.4 A neutral model of repertoire formation applied to DN3

In a recent article P. de Greef and colleagues [de Greef et al., 2020] developed a birth and death model for the diversity of the naive T cell pool under thymus input. This model can be readily applied to the DN3 pool. Under this model, we consider a population of cells of constant size N, any cell in the pool can die with rate $1/N$ and in parallel to compensate this loss receives an input from a source compartment (in our case DN3a) with a probability θ, or with a probability 1-τ, one cell from the pool is randomly chosen to divide. θ therefore encodes the relative contribution of proliferation and input to the maintenance of the compartment. It is a neutral model in the sense that each cell in the pool as an equal probability to divide with rate (1-τ)/N. Such a system will reach a steady state where the clone size distribution is constant, consequently, the number of clones of size 1 that are lost by death or differentiation Let F_i be number of clones of size i cells, so that F_1 is the number of distinct β chain clones of size one. At steady state, the number of clones of size 1 that are lost by death or differentiation , F_1/N must be equal to θ the number of clones which enter the pool upon β chain rearrangement, so that $F_1 = \theta N$. iF_i represents the total number of cells of the F_i clones of size i, at steady state the loss rate $\frac{iF_i}{N}$ must be compensated by division coming from clones of size $i-1$, so that $\frac{iF_1}{N} = \frac{(i-1)F_{i-1}(1-\theta)}{N}$. Consequently, the authors find that:

$$F_i = \frac{F_1(1-\theta)^i - 1}{i} \tag{3.24}$$

This true clone size distribution however is never available, and actual data is always victim of a sampling effect, as only a small fraction of the actual population is actually sequenced. The authors estimate the expected sampled clone size distribution as:

$$\widehat{F_i} = F_i(\frac{s}{s+(1-s)\theta})^i \tag{3.25}$$

Where s is the fraction sampled from the total cell population.

Shannon's entropy is sometimes used in order to quantify diversity of an immune repertoire [Mora and Walczak, 2018]. It is a measure of the degree of uncertainty associated with a random variable. With respect to immune repertoire analysis, it represents the uncertainty in the identity of a clone sampled from a repertoire. It is maximal when all clones have the same probability of being sampled, that is, when no clone is overrepresented in the population. and it is given as:

$$H = -\sum_{i=1}^{S} p_i log p_i \tag{3.26}$$

Where S is the total number of species, i; P_i the probability to sample an individual from species i. Here we define the entropy as $H/Hmax$ with $Hmax$ the maximum possible entropy of the data, where all i are equiprobable.

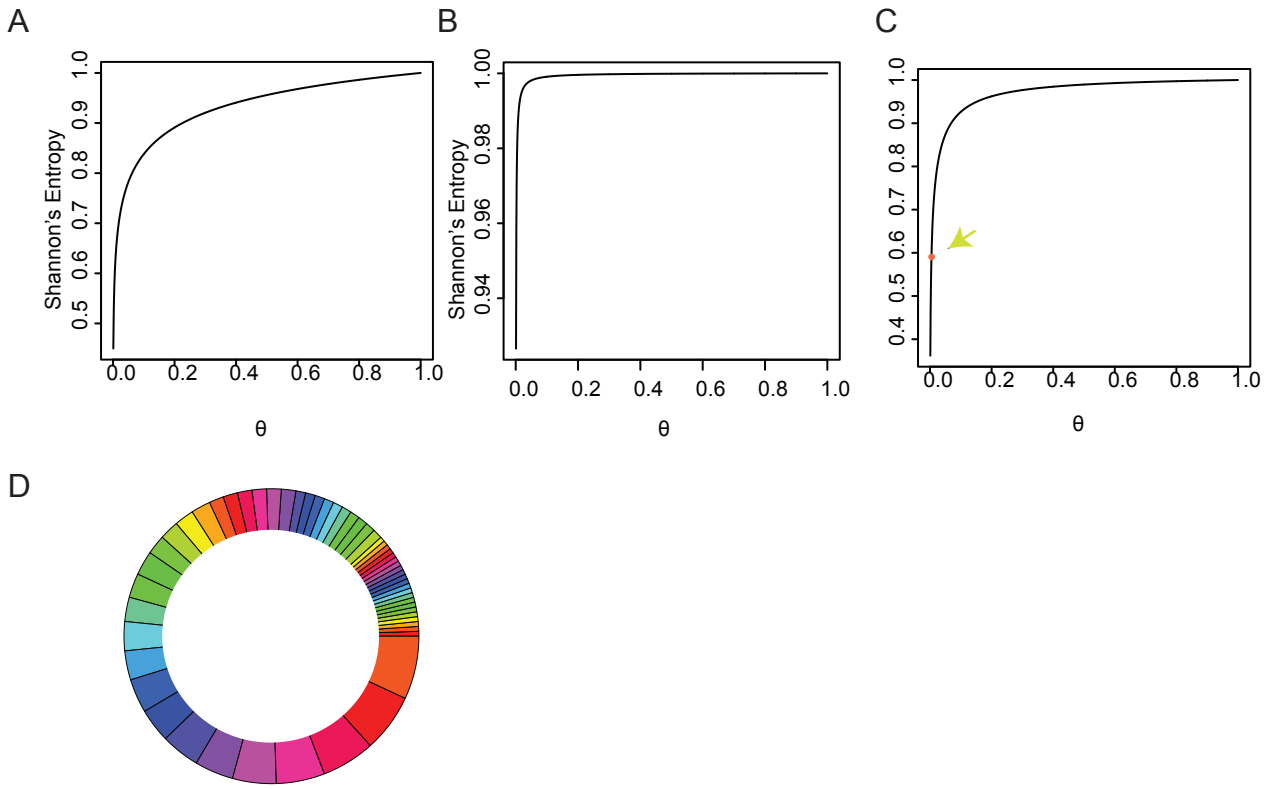


Figure 15: Results from the Neutral model of clonal diversity (A-C) shannon s entropy as a function of de novo β chain recombination θ (A) in a steady state DN3 b population (100 000 cells) (B) 0.1% fraction of cells sampled fom the general population (C) in 580 cells sampled from a population of 5800 cells (transplanted thymus) (D) expected clone size distribution for population sampled from the autonomous thymus

We can observe in Fig. 15 A that under the neutral model, clone diversity is directly dependent on θ. The diversity increases as the contribution of proliferation decreases relative to influx. In order to balance the chances for each β chain to be selected by positive and negative selection, an optimal strategy is to maximize diversity of the pre-selection repertoire. This suggests that the level of amplification of recombined repertoire should be tightly controled so as to avoid over-representation of specific clones.

The effect of sampling on the entropy is apparent in figure 15 B. Here we show infered the relation between θ and entropy in samples of size corresponding to our single cell RNA seq data in control mice. We observe that maximum entropy is already reached for very low input. Consequently, the observed β chain repertoire diversity in the DN3 population is not conclusive with respect to the actual dynamics of the DN3 population.

In transplanted thymi, the cellularity is in general 10 time lower than in control thymi and moreover, the DN3 population only represents 0.1% of transplanted thymocytes compared to around 1% of thymocytes in a control mouse. Consequently, while we sequence a similar number of cells in both cases, the sampled fraction is 100X higher for transplanted thymi than for control thymi.

Fig.15 C shows that, in this context, the clone size distribution can contain information relative to θ. The orange dot shows the entropy of the sample presented in Fig. 14. This relatively low entropy suggests a very low contribution of de novo to the steady state, with a θ of 0.005. According to this picture, the maintenance of the DN3b population is mainly ensured by proliferation and the preDN3b cells contribute very little to thymus autonomy. There is an expected clone size distribution associated with such a low θ which we present in figure 15 C. This predicted distribution differs significantly from the observed distribution, in Fig. 14 C. In the model, a large fraction of the clones have size superior than one but no clones have more than 20 cells. In the tranplanted thymus, the largest clones represents 25% of the cells and two other clones make up around 10% of the cells each but more importantly, 72% of the clones are only represented once. Such a strong discrepancy between the model and the data suggests that the neutral model is not adequate to describe the dynamics in autonomous DN3. There are no dozens of alternative to this neutral model, If we assume that the probability of generating twice the same TCR sequence de novo is negligible, then the only alternative to the model the we present here is a model where clones are not equal, either in terms of differentiation/death probability or proliferation probability. While it is known that the majority of the transplanted automomous thymi end up developing T cell acute lymphoblastic leukemia after 16 weeks and more , unpublished work from my colleague Verena Körber (personal communication) has established that the development of large mutant clones only occurs weeks after our experiment. It is now established that preTCR interaction with ligands has an effect on proliferation

3.3.5 Effect of thymus autonomy on $\gamma\delta$ T-cell differentiation

The model of DN3 differentiation that we presented above is a simplifaction which ignores an important dimension of T cell differentiation. At the DN3 stage, thymocyte are still bipotent, they can become $\alpha\beta$ or $\gamma\delta$ T cells. In adult mice, $\gamma\delta$ T cells make up 1 to 4% of T cells in the thymus and

lymphoid organ but they are very frequent in most mucosal sites, they make up 20 to 40% of the intestinal intraepithelial T cells of T cells [Chien et al., 2014]. They are critical to the maintenance of of a civil relationship between mammalian cells and its microbiome. Their TCR recognize in particlar native antigens on the surface of bacteria, or lipids presented by MHC like surface proteins.

Both γ and δ chains are recombined at the DN2-DN3 stages, and it is now generally accepted that TCR signaling determines the fate of these progenitors on a first come first serve basis. The pre-TCR (PTCRA+β chain) activation induces a weak signal leading to differentiation, while $\gamma\delta$ TCR activation produces a distinctively strong signal [Ciofani and Zuniga-Pflucker, 2010] sufficient to induce full maturation of the $\gamma\delta$ T cells.

The $\alpha\beta$ lineage is likely favored because a $\gamma\delta$ TCR requires two productive recombination events while the pre-TCR only requires one, the β chain. If, as we observe in the autonomous thymus, TCR beta activation does not induce systematic commitment to the $\alpha\beta$ lineage, then we can expect that DN3 cells which have passed the beta selection can still become $\gamma\delta$.

At this stage, we did not obtain enriched transcriptomes for γ and δ chains that we would need to reconstitute the $\gamma\delta$ repertoire in the DN3, but our full transcriptome data indicates the autonomous DN3 overexpress the gamma and delta chains compared to control. To study the potential change in the $\gamma\delta$ differentiation process, we have now sorted $\gamma\delta$ T cells from control or autonomous thymi as well as DN3 from the same mice and performed single-cell transcriptome+VDJ sequencing. This data should indicate whether a cell which has gone through the beta selection has still the potential to become $\alpha\beta$ lineage T cell under thymus autonomy.

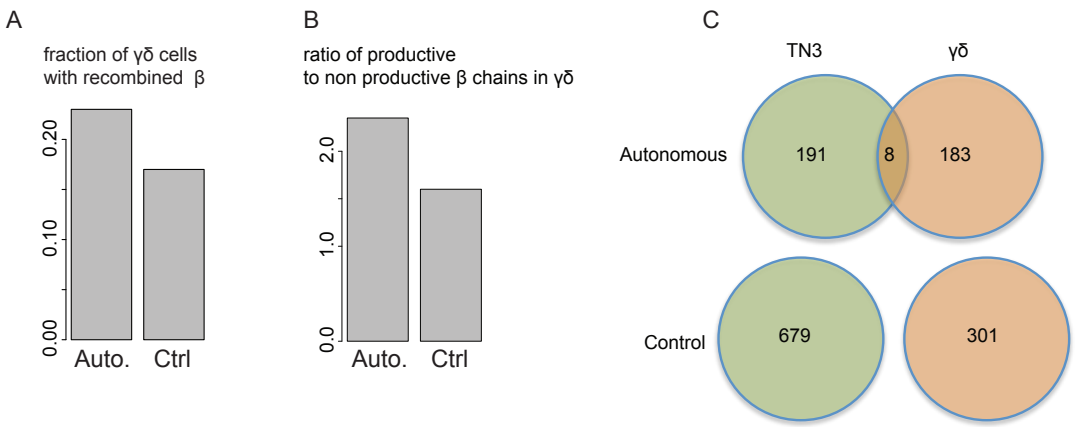


Figure 16: Overview of the β chain repertoire in γδ and DN3. (A) fraction of γδ cells with recombined β chain, (B)ratio of productive to non productive β chains (C) β repertoire diversity and overlap between SN3 and γδ

If γδ vs αβ branching occurs before the β selection, one can assume an absence of selection pressure for the β chain, and , consequently, when considering that β rearrangement can happend on both loci, only, at most 55% of γδ T cell should carry a productive β chain. Therefore, we expect a ratio of productive to non productive to be of 1.2 maximum. In both control and transplanted thymi (Fig. 16 A) more than 20% of γδ express a recombined β chain, among these cells. We observe in both cases a ratio of productive to non productive which is higher than the maxmimum expected value under the assumption that selection precludes γδ commitment. Furthermore, the ratio is higher in the transplanted sample than in the control: This strongly suggests that thymus autonomy favors the differentiation of TN3 in γδ beyond the beta selection.

The analysis of clones shows an overlap between TN3 and γδ T cells, in the transplanted thymus, this could mean that cells carrying a recombined β chain can function as progenitors for γδ T cells while remaining within the pool of CD25+ αβ T cell progenitors.

We do not observe any overlap in the control but, as mentioned previously, the sampling effect make it impossible make definitive statements on the structure of the repertoire in the control mice. This feature supports our hypothesis that influx influences γδ versus αβ commitment.

4 Discussion

4.1 Prediction of cell-cell interations from transcriptomic data

Inference of receptor-ligands interactions from transcriptomic data is a rapidly evolving topic, and others have been working on it while I was developing my own method. The main challenge to this endeavour is to define a metric to select valuable candidate interactions. Some authors have tried to infer receptor-ligand interactions based on the activity of signaling pathways downstream of the receptor [Choi et al., 2015, Browaeys et al., 2020]. The goal of such methods is to identify the ligands responsible for changes in the transcriptome of the target cells. This is a valid approach when one seeks to study a change of phenotype and identify the cause of the change or the difference between cell types. However, this approach does provide a full map of a cell-to-cell interactome as it only detects changes between conditions and will not identify unaffected interactions. Furthermore, this approach is limited by available knowledge of transcriptomic pathways, and it implies that the ligand-receptor interaction is followed by transcriptomic effect, which is not necessarily the case, for contact between cell adhesion molecules for instance.

Other methods published recently have focused only on expression of receptor and ligand such as CellPhoneDB [Efremova et al., 2020] which takes into accounts receptor complexes and select interactions based on co expression of the receptors components. The coexpresssion of the multiple receptor components increases the confidence in the interaction. The main advantage of our method is to look at ranks of receptors and ligands and not directly at expression values which allows the easy integration of different datasets. Furthermore, our interaction scoring system per-

mits the global comparison of cell to cell interactomes as well as the identification of specific interactions.

4.2 Contribution to the understanding of bone marrow niche

Many authors now consider that hematopoietic stem cells are heterogenous, and are made of several subpopulations with distinct functional properties and occupying spacially distinct niches in the bone marrow ([Pinho et al., 2018, Zhao et al., 2019]). These subtypes differ in proliferations, differentiation and and bone marrow reconsitution potential. Some authors have notably pointed out that HSCs from the endosteal region, differed from central bone marrow HSCS by bone marrow reconstitution potential while sharing the same phenotype [Grassinger et al., 2010]. Our analysis suggests that signals received by the HSPCs depends on their localisation. Indeed, different niche cell populations (Osteoblasts, MSCs and PaS) sorted based on distinct cell surface markers but occupying the endosteal region of the bone marrow tend to produce similar signals to hematopoietic cells, in particular we observe a clear endosteal vs central bone marrow effect. This differential regional effect could affect hematopoietic stem cells. Spatial transcriptomics has now become a hot topic and a paper published this year [Baccin et al., 2020] actually present a method to associate bone marrow niche cell transcriptomes to their fine-grained localization in the bone marrow. The authors do not characterize niche-HSPCs interactomes precisely and focus mainly on differential expression of Cxcl12 and Kitl by stromal cells to identify valuable niche cell candidates. However their spatial transcriptomic approach could constitute a valid way to study differential regulation of HSCs depending on the environment.

4.3 Dynamic regulation of hematopoietic cells by the bone marrow niche

Our analysis also revealed that cell population expressing specific cytokines often express their antagonist, albeit at a lower level. Such is the case for instance for Cxcl12 and its antagonist Cxcl14 expressed by MSC BM, or Jag1 and its antagonist Egfl7 expressed by endothelial cells. This feature could indicate that niche cells can modulate their effect on HSCs. Cxcl12 is a homing chemokine, it attracts the HSPCs to their niche and its expression by MSCs has already been shown to be controled by the sympathetic nervous system [Katayama et al., 2006, Mendez-Ferrer et al., 2008]. The production of Cxcl14, the Cxcl12 inhibitor could conversely ensure a fast release of the HSCs from their niche. The actual spatio-temporal expression of these antagonists and their role remain unknown and single cell RNA-seq could potentially reveal whether whether the same cells can indeed express both or whether distinct subpopulations of cells express one of the two agonists to the exclusion of the other.

Another aspect, of HSC-niche interactions which has raised some interests in the past few years is the regulation of niche cells by HSCs. So far, a few publications have tackled the question. The clearest evidence of such regulation was published by Sean Morrisson lab in 2015 [Zhou et al., 2015],the authors show that Angpt1 produced by HSPCs regulates niche regeneration after irradiation. A publication from the Höfer lab, to which I contributed modestly, envisaged on a theoretical view that HSCs-niche homeostasis depends on activation of niche cell proliferation by HSCs. This theoretical model could explain a series of independent observations, however no direct evidence of this has been published [Becker et al., 2019]. Our method could in principle be used to search for potential interactions between niche and HSPCs where the niche provides ligands to HSPCs and contribute to the appraisal of the hypothesis developed in this paper.

4.4 Estimation of cell proliferation based on short term thymidine analogue pulse-chase experiment

One of the strengths of our cell cycle method is that it evaluates explicitly some technical limitations of the experiment. One feature that we estimate in particular is the kinetic of EdU incorporation. Some papers published in the last few years which have used short term pulse-chase thymidine analogue to estimate cell proliferation make very specific assumptions regarding the kinetic degradation of EdU which can lead to important errors. This issue is particularly acute in the double staining experiment (EdU-BrdU) described by [Akinduro et al., 2018] and colleagues where the evaluation of entry into S phase relies one single time point. The authors label cells with 2 distinct thymidine analogues at 2h of interval. They assume that the first analogue (EdU) only labels cells which were in S phase at the time of injection as it is immediately degraded, consequently they consider that cells which have only incorporated the second analogue represent the entirety of the cells which have entered S phase during the 2 hours gaps. They use this value directly to compute the entry in S phase per hour. According to our own observations and model fitting, the peak of labeled cells in S phase is reached around the first hour after injection and the diminution of the EdU incorporation is gradual. This observation is clearly incompatible with the assumptions made by the authors. Although the biological system that they study is distinct from ours, this points to potential considerable error that can stem from such assumption.

We use total DNA content and thymidine analogue to estimate cell cycle progression. Classically, studies which have used this approach rely on assumptions relative to the content of a gate at a given time [White et al., 1990, Kretschmer et al., 2020]. Flow cytometry gates are set differently depending on the time point and assume that very specific subpopulation of cells will be in a gate at a particular time point. Typically at a given time point divided EdU+ cells are separated from undivided EdU positive cells. For this strategy to be reasonable, such a time window with perfectly separable divided and undivided EdU+ cells must exist one has to obtain a data point at a specific time where these subpopulation are separable. We do not need to make assumptions to the content of gates at a certain time, but only on what cells can be in a gate at any time, Our only assumptions is that the G01 gate will contains cells in G0/G1 and S, that the S gate will contain

cells in S phase and that the G2/M gate will contain cells in S and G2/M.

4.5 The difficult definition of G0 cells

The concept of cell quiescence arose with the first proliferation assays using radioactive nucleotide to label cells with replicating DNA [So and Cheung, 2018]. It was then observed that some cells do not incorporate the labels and those were then refered as "non-cycling". People then defined quiescence as a reversible state to distinguish it from terminally differentiated and senescent cells lacking the potential to return to the cell cycle. The molecular characterization of quiescent cells proved difficult and it is still often defined negatively by an absence of proliferation markers rather than identified by sets G0 specific of markers. It is generally observed that these cells are smaller, contain less RNA and are metabolically less active than proliferating cells. The entry into G0 remains also in debate. While early models proposed the existence of a restriction point within the G1 phase [Giacinti and Giordano, 2006] with the protein Rb maintaining cells in an arrested cells while non-phosphorylated, more recent works point to a decision made before the end of the previous cell cycle [Spencer et al., 2013] The lack of clear molecular marker make it difficult to identify these cells in vivo. The proliferation marker Ki67 is widely used to separate G0 cells from G1 by flow cytometry but recent work has established that it is not in fact a binary marker of proliferation versus quiescence [Miller et al., 2018]. Even current single-cell RNA seq technology which have improved of understanding of cell heterogeneity cannot identify quiescent cells. While some tools can distinguish S and G2M cells based on expression of specific markers [Butler et al., 2018], the G0-G1 compartment remains very much a black box. We demonstrate with DPs that our method can detect transition from proliferative state to a quiescent cells as EdU labeled cells do not transit to S phase. Cell entry into quiescence can now be detected in vitro by labeling of proliferation associated protein and time lapse microscopy([Zerjatke et al., 2017]), however, single cells cannot be tracked easily over long period of time in vivo and to our best knowledge, no method was able to identify transition from cycling to quiescent in vivo.

Quiescence is a hallmark of adult stem and progenitor cells, in particular in the bone marrow where the transition between states of quiescence and proliferation is generally thought to be con-

troled by the niche. While this question has been studied since at least seminal work of R.Schofield, there has been very little description of transition between a state of proliferation and quiescence in this context for every reasons that we just mentioned. Naturally, one of the perspective of our method would be to quantify this transition. Ann-Katrin Fanti from the group of Hans-Reimer Rodewald has generated data of the same kind as our own thymocyte data for HSPCs. However the cell cycle of early hematopoietic progenitor is significantly longer that that of thymocytes so that we could not observe reentry of EdU labeled cells into S phases within 18h. Another issue is the very low number of proliferating HSCs which increases the inter-sample variability. We might be able to remedy this issue with a second EdU pulse which would increase the frequency of labeled cells. Despite these difficulties, there are no fundamental preclusion to the use of our method to quantify transition from proliferation to quiescence in HSCs.

So far, our method works only in a context where input is negligible relative to proliferation as we assume that both steady state cell cycle distribution and kinetics of EdU labeled cells only depend on the proliferation dynamics of the compartment. There can be situations where the contribution of upstream progenitors is not negligible compared to proliferation of the reference compartment, in such cases, the influx of cells from the upstream compartment can affect the fraction of EdU labeled in the reference compartment. The description of EdU dynamics in the upstream compartment could potentially provide information on the input but the problem may become intractable if the source compartment is unknown.

Another limitation of our method is that we do not allow a return from G0, although quiescence is generally defined as a reversible state. The observation of EdU labeled cell returning to a cell cycle after an extended period of quiescence would probably require a longer experiment where EdU dilution upon cell division and input from upstream compartment would become significant factors to take into account.

4.6 Alterations of DN3 differentiation programme in the autonomous thymus

In this work we observed that in the autonomous thymus, the DN3 population has self-renewal capacity and seem to act as "thymic stem cells". unlike in control mice, the majority of DN3 thymocytes carried recombined β chains. β chains rearrangement however seemed to be going on as we could find cells expressing germline chains, a hallmark of rearrangement. If rearrangement still occurs, even at a reduced level it must signify that cells upstream of DN3b proliferate to maintain the pool of cells with "rearrangement potential". Interestingly, in the autonomous thymus, the cells rearranging the β chain could not be distinguished from cells carrying a β chain at a transcriptional level.

In a similar manner as the transcriptome, the cell cycle distribution (total DNA content) suggest that, while in control DN3, only cells with productive intra cellular β chain expression actually proliferates, proliferation occurs in the autonomous thymus irresepective of their level of intra cellular β chain expression. The quiescence of the DN3a is known to result of a desensitization to Il-7 signaling [Hosokawa and Rothenberg, 2018]. it could be that higher availability of Il-7 in absence of competitors overcome this desensitization.

Our work on T- cell development focused particularly on the DN3 already carrying a recombined β chain, the large majority of the DN3 in the autonomous thymus. The tight link between proliferation and differentiation at the stage of β selection described by [Kreslavsky et al., 2012] is now an accepted feature of T-cell development. Here we observe cells carrying recombined β chain capable of self sustaining when they do not have to compete for Il-7 with bone marrow progenitors. This result suggests that Il-7 can brake differentiation after the β selection. This finding nuances reports that proliferation and differentiation are tighly coupled at the β selection stage [Kreslavsky et al., 2012]. It seems however that this process does not affect T cell progenitors evenly as some clones are amplified beyond what would be expected under a neutral model where each progenitor has an equal chance to divide and differentiate. The question remain whether this non neutral effect stems from preTCR affinity for its ligands.

4.7 Il-7 as a determinant of $\gamma\delta$ versus $\alpha\beta$ commitment

We then use this mouse model to investigate the effect of such reduction of differentiation on $\gamma\delta$ T cell which diverge from the $\alpha\beta$ T cells at the DN3 stage. The identification of the g$\gamma\delta$ vs. $\alpha\beta$ branching point as long been debated. Two main conceptions have been opposed, pre-commitment and TCR signaling. According to the pre-commitment hypothesis, early T-cell progenitors are committed to one of the two lineages before they even express recombined TCR. A study has notably shown that Il7R expression at the DN2 stage was associated with a strong bias to the $\gamma\delta$ lineage while no recombined TCR $\gamma\delta$ is expressed before the DN3 stage [Kang et al., 2001]. The authors of this study also observed that the higher Il-7 receptor expression did not affect the $\gamma\delta$ recombination, thereby excluding an effect on the onset of TCR activation.

The other hypothesis supports that preTCR and $\gamma\delta$ signals are different in nature and that cells are instructed into one of the two lineages based on the TCR signal they receive. This second hypothesis became generally accepted following observations that the three $\gamma\delta$ and β chains recombine simultaneously [Kreslavsky et al., 2010] and that TCR signal strength determines whether cells differentiate into one lineage or the other, with a weak signal inducing commitment to the $\alpha\beta$ lineage [Hayes et al., 2005, Haks et al., 2005].

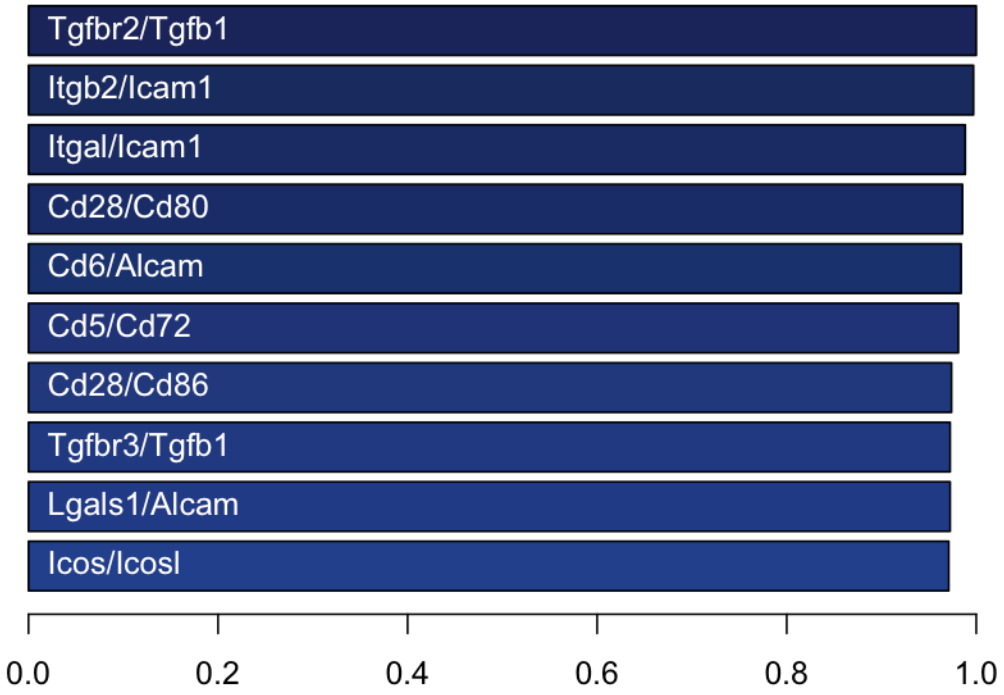
While the second proposition is now widely accepted, it could so far not be reconciled with the first observation that Il-7 receptor expression before TCR rearrangement correlates with a very strong fate bias. In this work, we show that in a context of low competition for Il-7 in the autonomous thymus, DN3 cells carrying productive β chain do not immediately differentiate into DN4 and can continuously function as a source of progenitors. This decreased differentiation could be explained by a simple mathematical model where the excess of a resource, in this case Il-7, decreases differentiation. This conclusion is supported by in vitro experiments showing that high concentrations of Il-7 reduces DN3 differentiation into DP.

Concurently, we observe an increase in the ratio of productive to non productive β chain rearrangement in $\gamma\delta$ T cells, suggesting that a significant fraction of these cells are β-selected under thymus autonomy. Taken together, these elements point to a new role for Il-7 as a modulator of $\gamma\delta$ versus $\alpha\beta$ differentiation by delaying $\alpha\beta$ commitment following β selection. This would explain

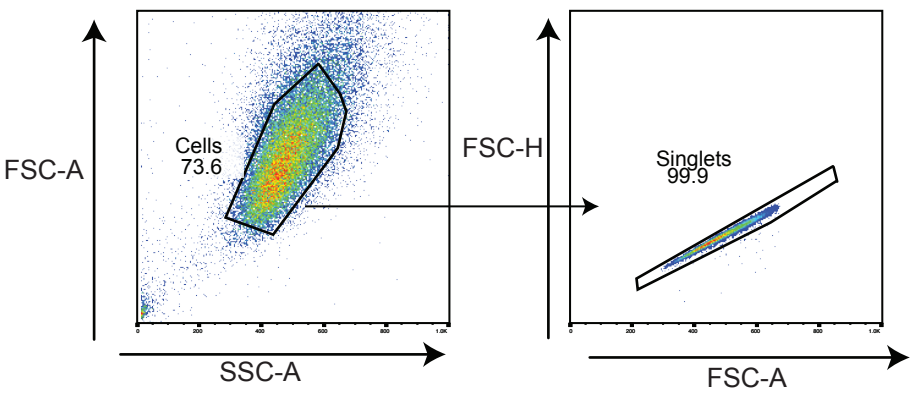
why DN2 with higher Il-7 expression are more prone to become $\gamma\delta$, this bias would not stem from any pre-commitement but from a slower differentiation of these cells.

In order to support a direct effect of Il-7 on thymus autonomy, we could also assay the activation of the pathway in the DN3 in control and in autonomous thymi. At this stage, Stat5 phosphorylation, which can be assayed by flow cytomety is often used as a proxy for Il-7 activation. Since Il-7 production by TEC can be dynamically regulated in vivo [Alves et al., 2009], we could also evaluate il-7 production in control vs. autonomous thymus. More generally, RNAsequencing of TEC could provide information on changes incurred by TEC in thymus autonomy.

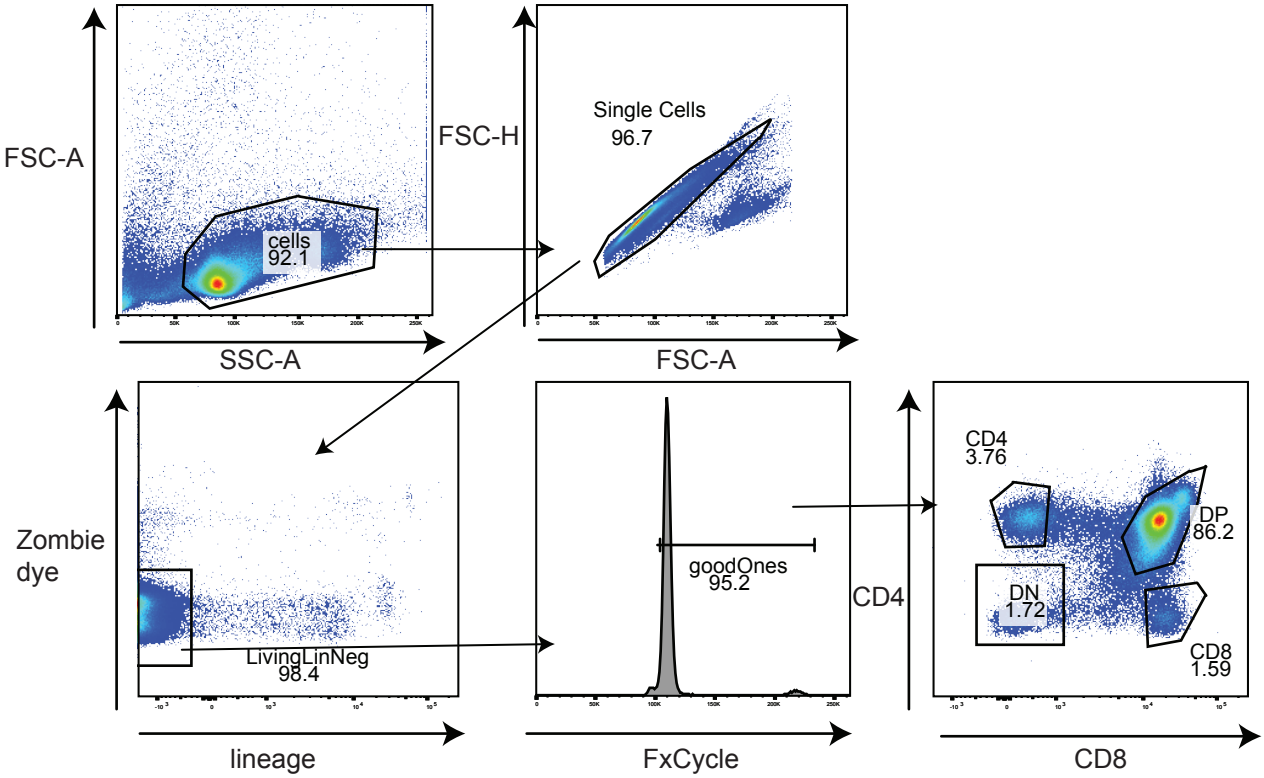
5 Supplemental Figures



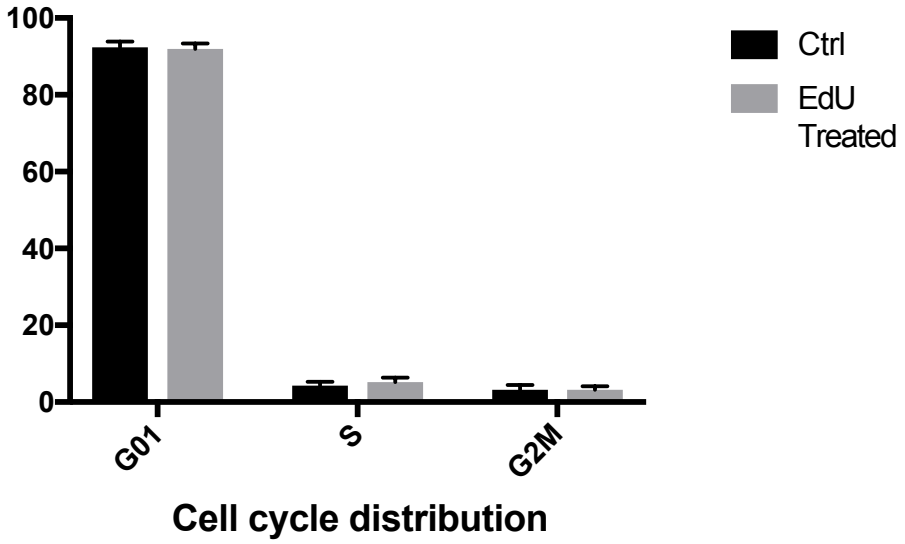
Sup. Figure 1: highest interactions scores between murine CD4+ Dendritic cells and CD4+ naive T cells



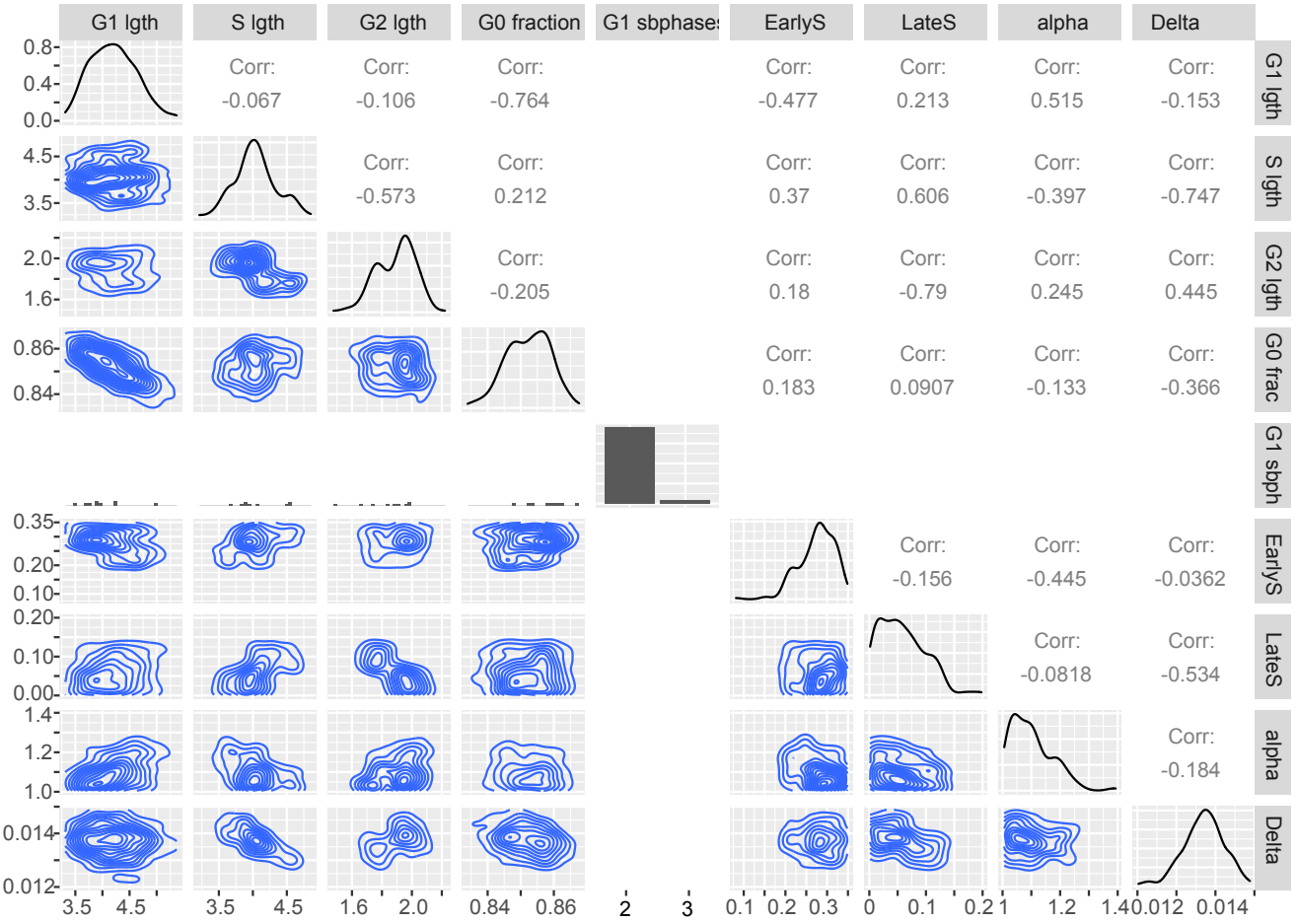
Sup. Figure 2: Gating stragety Tet21N. Cells are first gated based on size and doublets are eliminated based on the forward scater (area vs. height).



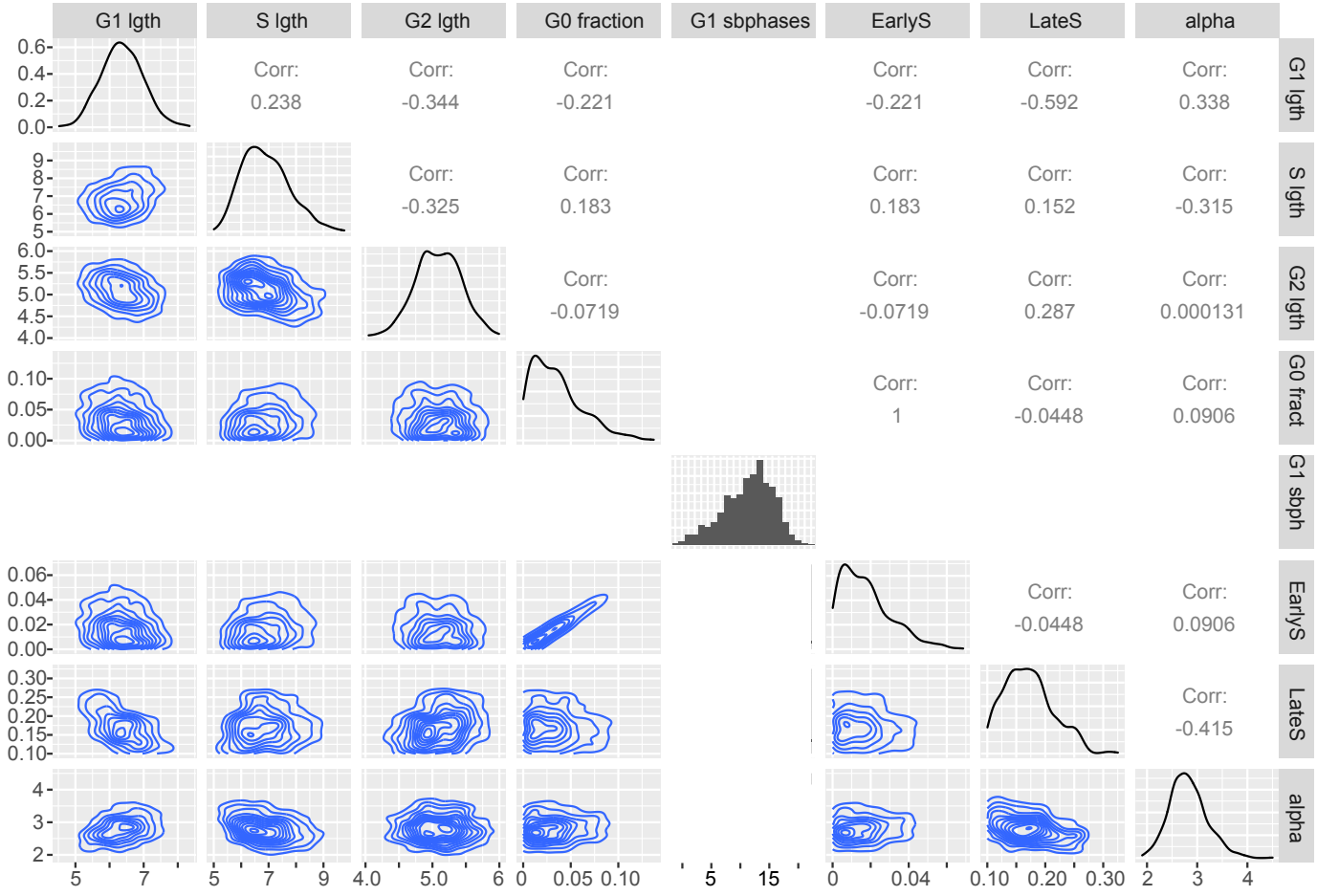
Sup. Figure 3: Gating stragety for thymocytes. Cells are first gated based on size, doublets are eliminated based on the forward scater (area vs. height), then dead cells and lineage positive cells are excluded and finally we exclude cells with abnormal DNA content.



Sup. Figure 4: Cell cycle distribution in control mouse or in EdU treated mouse (time points 3h,5 an 12h were pooled). T-test revealed no significant changes between conditions



Sup. Figure 5: posterior distribution for cell cycle parameters in Double positive Thymocytes and correlations coefficients between parameters(Pearson). G1 length, S length anf G2 length are expressed in hours, EarlyS and LateS are expressed as fractions of the total S phase, EdU incorporation rate alpha and death delta are rates per hour



Sup. Figure 6: posterior distribution for cell cycle parameters for Tet21N and correlations coefficients between parameters(Pearson). G1 length, S length anf G2 length are expressed in hours, EarlyS and LateS are expressed as fractions of the total S phase, EdU incorporation rate alpha

Bibliography

[Akinduro et al., 2018] Akinduro, O., Weber, T. S., Ang, H., Haltalli, M. L. R., Ruivo, N., Duarte, D., Rashidi, N. M., Hawkins, E. D., Duffy, K. R. and Lo Celso, C. (2018). Proliferation dynamics of acute myeloid leukaemia and haematopoietic progenitors competing for bone marrow space. Nat Commun *9*, 519.

[Alves et al., 2009] Alves, N. L., Richard-Le Goff, O., Huntington, N. D., Sousa, A. P., Ribeiro, V. S., Bordack, A., Vives, F. L., Peduto, L., Chidgey, A., Cumano, A., Boyd, R., Eberl, G. and Di Santo, J. P. (2009). Characterization of the thymic IL-7 niche in vivo. Proc. Natl. Acad. Sci. U.S.A. *106*, 1512–1517.

[Axelrod and Pienta, 2014] Axelrod, H. and Pienta, K. J. (2014). Axl as a mediator of cellular growth and survival. Oncotarget *5*, 8818–8852.

[Baccin et al., 2020] Baccin, C., Al-Sabah, J., Velten, L., Helbling, P. M., Gruenschlaeger, F., Hernandez-Malmierca, P., Nombela-Arrieta, C., Steinmetz, L. M., Trumpp, A. and Haas, S. (2020). Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. Nat. Cell Biol. *22*, 38–48.

[Baron and Penit, 1990] Baron, C. and Penit, C. (1990). Study of the thymocyte cell cycle by bivariate analysis of incorporated bromodeoxyuridine and DNA content. Eur. J. Immunol. *20*, 1231–1236.

[Becker et al., 2019] Becker, N. B., Guenther, M., Li, C., Jolly, A. and Hoefer, T. (2019). Stem cell homeostasis by integral feedback through the niche. J. Theor. Biol. *481*, 100–109.

[Boehm, 2012] Boehm, T. (2012). Self-renewal of thymocytes in the absence of competitive precursor replenishment. J. Exp. Med. *209*, 1397–1400.

[Boudil et al., 2015] Boudil, A., Matei, I. R., Shih, H. Y., Bogdanoski, G., Yuan, J. S., Chang, S. G., Montpellier, B., Kowalski, P. E., Voisin, V., Bashir, S., Bader, G. D., Krangel, M. S. and Guidos, C. J. (2015). IL-7 coordinates proliferation, differentiation and Tcra recombination during thymocyte Îš-selection. Nat. Immunol. *16*, 397–405.

[Browaeys et al., 2020] Browaeys, R., Saelens, W. and Saeys, Y. (2020). NicheNet: modeling intercellular communication by linking ligands to target genes. Nat. Methods *17*, 159–162.

[Busch et al., 2015] Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S. M., Reth, M., Hoefer, T. and Rodewald, H. R. (2015). Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. Nature *518*, 542–546.

[Butler et al., 2018] Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

[Calvi et al., 2003] Calvi, L. M., Adams, G. B., Weibrecht, K. W., Weber, J. M., Olson, D. P., Knight, M. C., Martin, R. P., Schipani, E., Divieti, P., Bringhurst, F. R., Milner, L. A., Kronenberg, H. M. and Scadden, D. T. (2003). Osteoblastic cells regulate the haematopoietic stem cell niche. Nature *425*, 841–846.

[Cavanagh et al., 2011] Cavanagh, B. L., Walker, T., Norazit, A. and Meedeniya, A. C. (2011). Thymidine analogues for tracking DNA synthesis. Molecules *16*, 7980–7993.

[Chao et al., 2019] Chao, H. X., Fakhreddin, R. I., Shimerov, H. K., Kedziora, K. M., Kumar, R. J., Perez, J., Limas, J. C., Grant, G. D., Cook, J. G., Gupta, G. P. and Purvis, J. E. (2019). Evidence that the human cell cycle is a series of uncoupled, memoryless phases. Mol. Syst. Biol. *15*, e8604.

[Chatr-aryamontri et al., 2007] Chatr-aryamontri, A., Ceol, A., Palazzi, L. M., Nardelli, G., Schneider, M. V., Castagnoli, L. and Cesareni, G. (2007). MINT: the Molecular INTeraction database. Nucleic Acids Res. *35*, D572–574.

[Cheraghali et al., 1994] Cheraghali, A. M., Knaus, E. E. and Wiebe, L. I. (1994). Bioavailability and pharmacokinetic parameters for 5-ethyl-2'-deoxyuridine. Antiviral Res. *25*, 259–267.

[Chien et al., 2014] Chien, Y. H., Meyer, C. and Bonneville, M. (2014). ÎşÎť T cells: first line of de-

fense and beyond. Annu. Rev. Immunol. *32*, 121–155.

[Choi et al., 2015] Choi, H., Sheng, J., Gao, D., Li, F., Durrans, A., Ryu, S., Lee, S. B., Narula, N., Rafii, S., Elemento, O., Altorki, N. K., Wong, S. T. and Mittal, V. (2015). Transcriptome analysis of individual stromal cell populations identifies stroma-tumor crosstalk in mouse lung cancer model. Cell Rep *10*, 1187–1201.

[Ciofani and Zuniga-Pflucker, 2010] Ciofani, M. and Zuniga-Pflucker, J. C. (2010). Determining gamma delta versus alpha beta T cell development. Nat. Rev. Immunol. *10*, 657–663.

[Crane et al., 2017] Crane, G. M., Jeffery, E. and Morrison, S. J. (2017). Adult haematopoietic stem cell niches. Nat. Rev. Immunol. *17*, 573–590.

[de Greef et al., 2020] de Greef, P. C., Oakes, T., Gerritsen, B., Ismail, M., Heather, J. M., Hermsen, R., Chain, B. and de Boer, R. J. (2020). The naive T-cell receptor repertoire has an extremely broad distribution of clone sizes. Elife *9*.

[DeBoer, 2020] DeBoer, R. (2020). Biological Modeling of Populations. Utrecht University.

[DeBoer and Perelson, 2013] DeBoer, R. J. and Perelson, A. S. (2013). Quantifying T lymphocyte turnover. J. Theor. Biol. *327*, 45–87.

[Ding et al., 2012] Ding, L., Saunders, T. L., Enikolopov, G. and Morrison, S. J. (2012). Endothelial and perivascular cells maintain haematopoietic stem cells. Nature *481*, 457–462.

[Efremova et al., 2020] Efremova, M., Vento-Tormo, M., Teichmann, S. A. and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc *15*, 1484–1506.

[El Kassar et al., 2004] El Kassar, N., Lucas, P. J., Klug, D. B., Zamisch, M., Merchant, M., Bare, C. V., Choudhury, B., Sharrow, S. O., Richie, E., Mackall, C. L. and Gress, R. E. (2004). A dose effect of IL-7 on thymocyte development. Blood *104*, 1419–1427.

[Foreman-Mackey et al., 2013] Foreman-Mackey, D., Hogg, D. W., Lang, D. and Goodman, J. (2013). emcee: The MCMC Hammer. astro-ph.IM *125*, 306–312.

[Frey et al., 1992] Frey, J. R., Ernst, B., Surh, C. D. and Sprent, J. (1992). Thymus-grafted SCID mice show transient thymopoiesis and limited depletion of V beta 11+ T cells. J. Exp. Med. *175*, 1067–

1071.

[Gegonne et al., 2018] Gegonne, A., Chen, Q. R., Dey, A., Etzensperger, R., Tai, X., Singer, A., Meerzaman, D., Ozato, K. and Singer, D. S. (2018). Immature CD8 Single-Positive Thymocytes Are a Molecularly Distinct Subpopulation, Selectively Dependent on BRD4 for Their Differentiation. Cell Rep *24*, 117–129.

[Giacinti and Giordano, 2006] Giacinti, C. and Giordano, A. (2006). RB and cell cycle progression. Oncogene *25*, 5220–5227.

[Grassinger et al., 2010] Grassinger, J., Haylock, D. N., Williams, B., Olsen, G. H. and Nilsson, S. K. (2010). Phenotypically identical hemopoietic stem cells isolated from different regions of bone marrow have different biologic potential. Blood *116*, 3185–3196.

[Haks et al., 2005] Haks, M. C., Lefebvre, J. M., Lauritsen, J. P., Carleton, M., Rhodes, M., Miyazaki, T., Kappes, D. J. and Wiest, D. L. (2005). Attenuation of gammadeltaTCR signaling efficiently diverts thymocytes to the alphabeta lineage. Immunity *22*, 595–606.

[Hara and Tanegashima, 2014] Hara, T. and Tanegashima, K. (2014). CXCL14 antagonizes the CXCL12-CXCR4 signaling axis. Biomol Concepts *5*, 167–173.

[Hayes et al., 2005] Hayes, S. M., Li, L. and Love, P. E. (2005). TCR signal strength influences alphabeta/gammadelta lineage fate. Immunity *22*, 583–593.

[Heng et al., 2008] Heng, T. S., Painter, M. W., Elpek, K., Lukacs-Kornek, V., Mauermann, N., Turley, S. J., Koller, D., Kim, F. S., Wagers, A. J., Asinovski, N., Davis, S., Fassett, M., Feuerer, M., Gray, D. H., Haxhinasto, S., Hill, J. A., Hyatt, G., Laplace, C., Leatherbee, K., Mathis, D., Benoist, C., Jianu, R., Laidlaw, D. H., Best, J. A., Knell, J., Goldrath, A. W., Jarjoura, J., Sun, J. C., Zhu, Y., Lanier, L. L., Ergun, A., Li, Z., Collins, J. J., Shinton, S. A., Hardy, R. R., Friedline, R., Sylvia, K. and Kang, J. (2008). The Immunological Genome Project: networks of gene expression in immune cells. Nat. Immunol. *9*, 1091–1094.

[Hosokawa and Rothenberg, 2018] Hosokawa, H. and Rothenberg, E. V. (2018). Cytokines, Transcription Factors, and the Initiation of T-Cell Development. Cold Spring Harb Perspect Biol *10*.

[Hu et al., 2015] Hu, Z., Lancaster, J. N. and Ehrlich, L. I. (2015). The Contribution of Chemokines and Migration to the Induction of Central Tolerance in the Thymus. Front Immunol *6*, 398.

[Humphries et al., 2006] Humphries, J. D., Byron, A. and Humphries, M. J. (2006). Integrin ligands at a glance. J. Cell. Sci. *119*, 3901–3903.

[Kang et al., 2001] Kang, J., Volkmann, A. and Raulet, D. H. (2001). Evidence that gammadelta versus alphabeta T cell fate determination is initiated independently of T cell receptor signaling. J. Exp. Med. *193*, 689–698.

[Katayama et al., 2006] Katayama, Y., Battista, M., Kao, W. M., Hidalgo, A., Peired, A. J., Thomas, S. A. and Frenette, P. S. (2006). Signals from the sympathetic nervous system regulate hematopoietic stem cell egress from bone marrow. Cell *124*, 407–421.

[Kreslavsky et al., 2012] Kreslavsky, T., Gleimer, M., Miyazaki, M., Choi, Y., Gagnon, E., Murre, C., Sicinski, P. and von Boehmer, H. (2012). Beta-selection-induced proliferation is required for AlphaBeta T cell differentiation. Immunity *37*, 840–853.

[Kreslavsky et al., 2010] Kreslavsky, T., Gleimer, M. and von Boehmer, H. (2010). Alphabeta versus gammadelta lineage choice at the first TCR-controlled checkpoint. Curr. Opin. Immunol. *22*, 185–192.

[Kretschmer et al., 2020] Kretschmer, L., Flossdorf, M., Mir, J., Cho, Y. L., Plambeck, M., Treise, I., Toska, A., Heinzel, S., Schiemann, M., Busch, D. H. and Buchholz, V. R. (2020). Differential expansion of T central memory precursor and effector subsets is regulated by division speed. Nat Commun *11*, 113.

[Krueger et al., 2017] Krueger, A., Zi?tara, N. and ?yszkiewicz, M. (2017). T Cell Development by the Numbers. Trends Immunol. *38*, 128–139.

[Kumar and Geiger, 2017] Kumar, S. and Geiger, H. (2017). HSC Niche Biology and HSC Expansion Ex Vivo. Trends Mol Med *23*, 799–819.

[Kunisaki et al., 2013] Kunisaki, Y., Bruns, I., Scheiermann, C., Ahmed, J., Pinho, S., Zhang, D., Mizoguchi, T., Wei, Q., Lucas, D., Ito, K., Mar, J. C., Bergman, A. and Frenette, P. S. (2013). Arteriolar niches maintain haematopoietic stem cell quiescence. Nature *502*, 637–643.

[Lacombe et al., 1988] Lacombe, F., Belloc, F., Bernard, P. and Boisseau, M. R. (1988). Evaluation of four methods of DNA distribution data analysis based on bromodeoxyuridine/DNA bivariate data. Cytometry *9*, 245–253.

[Leif et al., 2004] Leif, R. C., Stein, J. H. and Zucker, R. M. (2004). A short history of the initial application of anti-5-BrdU to the detection and measurement of S phase. Cytometry A *58*, 45–52.

[Liu et al., 2010] Liu, P., Li, P. and Burke, S. (2010). Critical roles of Bcl11b in T-cell development and maintenance of T-cell identity. Immunol. Rev. *238*, 138–149.

[Lun et al., 2016] Lun, A. T., McCarthy, D. J. and Marioni, J. C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Res *5*, 2122.

[Lynn et al., 2008] Lynn, D. J., Winsor, G. L., Chan, C., Richard, N., Laird, M. R., Barsky, A., Gardy, J. L., Roche, F. M., Chan, T. H., Shah, N., Lo, R., Naseer, M., Que, J., Yau, M., Acab, M., Tulpan, D., Whiteside, M. D., Chikatamarla, A., Mah, B., Munzner, T., Hokamp, K., Hancock, R. E. and Brinkman, F. S. (2008). InnateDB: facilitating systems-level analyses of the mammalian innate immune response. Mol. Syst. Biol. *4*, 218.

[Martins et al., 2014] Martins, V. C., Busch, K., Juraeva, D., Blum, C., Ludwig, C., Rasche, V., Lasitschka, F., Mastitsky, S. E., Brors, B., Hielscher, T., Fehling, H. J. and Rodewald, H. R. (2014). Cell competition is a tumour suppressor mechanism in the thymus. Nature *509*, 465–470.

[Martins et al., 2012] Martins, V. C., Ruggiero, E., Schlenner, S. M., Madan, V., Schmidt, M., Fink, P. J., von Kalle, C. and Rodewald, H. R. (2012). Thymus-autonomous T cell development in the absence of progenitor import. J. Exp. Med. *209*, 1409–1417.

[Masamoto et al., 2017] Masamoto, Y., Arai, S., Sato, T., Kubota, N., Takamoto, I., Kadowaki, T. and Kurokawa, M. (2017). Adiponectin Enhances Quiescence Exit of Murine Hematopoietic Stem Cells and Hematopoietic Recovery Through mTORC1 Potentiation. Stem Cells *35*, 1835–1848.

[Mende et al., 2019] Mende, N., Jolly, A., Percin, G. I., G?nther, M., Rostovskaya, M., Krishnan, S. M., Oostendorp, R. A. J., Dahl, A., Anastassiadis, K., H?fer, T. and Waskow, C. (2019). Prospective isolation of nonhematopoietic cells of the niche and their differential molecular interac-

tions with HSCs. Blood *134*, 1214–1226.

[Mendez-Ferrer et al., 2008] Mendez-Ferrer, S., Lucas, D., Battista, M. and Frenette, P. S. (2008). Haematopoietic stem cell release is regulated by circadian oscillations. Nature *452*, 442–447.

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

[Michie and Zuñiga-Pflücker, 2002] Michie, A. M. and Zuñiga-Pflücker, J. C. (2002). Regulation of thymocyte differentiation: pre-TCR signals and beta-selection. Semin. Immunol. *14*, 311–323.

[Miller et al., 2018] Miller, I., Min, M., Yang, C., Tian, C., Gookin, S., Carter, D. and Spencer, S. L. (2018). Ki67 is a Graded Rather than a Binary Marker of Proliferation versus Quiescence. Cell Rep *24*, 1105–1112.

[Miller, 1961] Miller, J. F. (1961). Immunological function of the thymus. Lancet *2*, 748–749.

[Miller and Mitchell, 1967] Miller, J. F. and Mitchell, G. F. (1967). The thymus and the precursors of antigen reactive cells. Nature *216*, 659–663.

[Mingueneau et al., 2013] Mingueneau, M., Kreslavsky, T., Gray, D., Heng, T., Cruse, R., Ericson, J., Bendall, S., Spitzer, M. H., Nolan, G. P., Kobayashi, K., von Boehmer, H., Mathis, D., Benoist, C., Best, A. J., Knell, J., Goldrath, A., Joic, V., Koller, D., Shay, T., Regev, A., Cohen, N., Brennan, P., Brenner, M., Kim, F., Nageswara Rao, T., Wagers, A., Heng, T., Ericson, J., Rothamel, K., Ortiz-Lopez, A., Mathis, D., Benoist, C., Bezman, N. A., Sun, J. C., Min-Oo, G., Kim, C. C., Lanier, L. L., Miller, J., Brown, B., Merad, M., Gautier, E. L., Jakubzick, C., Randolph, G. J., Monach, P., Blair, D. A., Dustin, M. L., Shinton, S. A., Hardy, R. R., Laidlaw, D., Collins, J., Gazit, R., Rossi, D. J., Malhotra, N., Sylvia, K., Kang, J., Kreslavsky, T., Fletcher, A., Elpek, K., Bellemare-Pelletier, A., Malhotra, D. and Turley, S. (2013). The transcriptional landscape of alpha beta T cell differentiation. Nat. Immunol. *14*, 619–632.

[Mizoguchi et al., 2014] Mizoguchi, T., Pinho, S., Ahmed, J., Kunisaki, Y., Hanoun, M., Mendelson, A., Ono, N., Kronenberg, H. M. and Frenette, P. S. (2014). Osterix marks distinct waves of

primitive and definitive stromal progenitors during bone marrow development. Dev. Cell *29*, 340–349.

[Mora and Walczak, 2018] Mora, T. and Walczak, A. M. (2018). Quantifying lymphocyte receptor diversity. Systems Immunology , 183198.

[Morales-Mantilla and King, 2018] Morales-Mantilla, D. E. and King, K. Y. (2018). The Role of Interferon-Gamma in Hematopoietic Stem Cell Development, Homeostasis, and Disease. Curr Stem Cell Rep *4*, 264–271.

[Morikawa et al., 2009] Morikawa, S., Mabuchi, Y., Kuboty, Y., Nagai, Y., Niibe, K., Hiratsu, E., Suzuki, S., Miyauchi-Hara, C., Nagoshi, N., Sunabori, T., Shimmura, S., Miyawaki, A. Nakagawa, T., Suda, T., Okano, H. and Matsuzaki, Y. (2009). Prospective identification isolation, and systematic transplantation of multipotent mesenchymal stem cells in murine bone marrow. J. Exp. Med *11*, 2483–2496.

[Morrison and Kimble, 2006] Morrison, S. J. and Kimble, J. (2006). Asymmetric and symmetric stem-cell divisions in development and cancer. Nature *441*, 1068–1074.

[Ogawa et al., 1991] Ogawa, M., Matsuzaki, Y., Nishikawa, S., Hayashi, S., Kunisada, T., Sudo, T., Kina, T., Nakauchi, H. and Nishikawa, S. (1991). Expression and function of c-kit in hemopoietic progenitor cells. J. Exp. Med. *174*, 63–71.

[Orchard et al., 2014] Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., Campbell, N. H., Chavali, G., Chen, C., del Toro, N., Duesbury, M., Dumousseau, M., Galeota, E., Hinz, U., Iannuccelli, M., Jagannathan, S., Jimenez, R., Khadake, J., Lagreid, A., Licata, L., Lovering, R. C., Meldal, B., Melidoni, A. N., Milagros, M., Peluso, D., Perfetto, L., Porras, P., Raghunath, A., Ricard-Blum, S., Roechert, B., Stutz, A., Tognolli, M., van Roey, K., Cesareni, G. and Hermjakob, H. (2014). The MIntAct project–IntAct as a common curation platform for 11 molecular interaction databases. Nucleic Acids Res. *42*, D358–363.

[Parrish-Novak et al., 2002] Parrish-Novak, J., Foster, D. C., Holly, R. D. and Clegg, C. H. (2002). Interleukin-21 and the IL-21 receptor: novel effectors of NK and T cell responses. J. Leukoc. Biol. *72*, 856–863.

[Peaudecerf et al., 2012] Peaudecerf, L., Lemos, S., Galgano, A., Krenn, G., Vasseur, F., Di Santo, J. P., Ezine, S. and Rocha, B. (2012). Thymocytes may persist and differentiate without any input from bone marrow progenitors. J. Exp. Med. *209*, 1401–1408.

[Pinho and Frenette, 2019] Pinho, S. and Frenette, P. S. (2019). Haematopoietic stem cell activity and interactions with the niche. Nat. Rev. Mol. Cell Biol. *20*, 303–320.

[Pinho et al., 2018] Pinho, S., Marchand, T., Yang, E., Wei, Q., Nerlov, C. and Frenette, P. S. (2018). Lineage-Biased Hematopoietic Stem Cells Are Regulated by Distinct Niches. Dev. Cell *44*, 634–641.

[Puthier et al., 2004] Puthier, D., Joly, F., Irla, M., Saade, M., Victorero, G., Loriod, B. and Nguyen, C. (2004). A general survey of thymocyte differentiation by transcriptional analysis of knockout mouse models. J. Immunol. *173*, 6109–6118.

[Ramilowski et al., 2015] Ramilowski, J. A., Goldberg, T., Harshbarger, J., Kloppmann, E., Kloppman, E., Lizio, M., Satagopam, V. P., Itoh, M., Kawaji, H., Carninci, P., Rost, B. and Forrest, A. R. (2015). A draft network of ligand-receptor-mediated multicellular signalling in human. Nat Commun *6*, 7866.

[Rodda and McMahon, 2006] Rodda, S. J. and McMahon, A. P. (2006). Distinct roles for Hedgehog and canonical Wnt signaling in specification, differentiation and maintenance of osteoblast progenitors. Development *133*, 3231–3244.

[Rodewald et al., 1995] Rodewald, H. R., Kretzschmar, K., Swat, W. and Takeda, S. (1995). Intrathymically expressed c-kit ligand (stem cell factor) is a major factor driving expansion of very immature thymocytes in vivo. Immunity *3*, 313–319.

[Roozendaal and Mebius, 2011] Roozendaal, R. and Mebius, R. E. (2011). Stromal cell-immune cell interactions. Annu. Rev. Immunol. *29*, 23–43.

[Ryl et al., 2017] Ryl, T., Kuchen, E. E., Bell, E., Shao, C., Florez, A. F., Moenke, G., Gogolin, S., Friedrich, M., Lamprecht, F., Westermann, F. and Hoefer, T. (2017). Cell-Cycle Position of Single MYC-Driven Cancer Cells Dictates Their Susceptibility to a Chemotherapeutic Drug. Cell Syst *5*, 237–250.

[Sakaue-Sawano et al., 2008] Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H., Imamura, T., Ogawa, M., Masai, H. and Miyawaki, A. (2008). Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. Cell *132*, 487–498.

[Salic and Mitchison, 2008] Salic, A. and Mitchison, T. J. (2008). A chemical method for fast and sensitive detection of DNA synthesis in vivo. Proc. Natl. Acad. Sci. U.S.A. *105*, 2415–2420.

[Schmidt et al., 2009] Schmidt, M. H. H., Bicker, F., Nikolic, I., Meister, J., Babuke, T., Picuric, S., Mueller-Esterl, W., Plate, K. H. and Dikic, I. (2009). Epidermal growth factor-like domain 7 (EGFL7) modulates Notch signalling and affects neural stem cell renewal. Nat. Cell Biol. *11*, 873–880.

[Schofield, 1978] Schofield, R. (1978). The relationship between the spleen colony-forming cell and the haemopoietic stem cell. Blood Cells *4*, 7–25.

[Shannon et al., 2003] Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

[Sleckman et al., 1996] Sleckman, B. P., Gorman, J. R. and Alt, F. W. (1996). Accessibility control of antigen-receptor variable-region gene assembly: role of cis-acting elements. Annu. Rev. Immunol. *14*, 459–481.

[So and Cheung, 2018] So, W. K. and Cheung, T. H. (2018). Molecular Regulation of Cellular Quiescence: A Perspective from Adult Stem Cells and Its Niches. Methods Mol. Biol. *1686*, 1–25.

[Spencer et al., 2013] Spencer, S. L., Cappell, S. D., Tsai, F. C., Overton, K. W., Wang, C. L. and Meyer, T. (2013). The proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. Cell *155*, 369–383.

[Stark et al., 2006] Stark, C., Breitkreutz, B. J., Reguly, T., Boucher, L., Breitkreutz, A. and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. Nucleic Acids Res. *34*, D535–539.

[Sugiyama et al., 2006] Sugiyama, T., Kohara, H., Noda, M. and Nagasawa, T. (2006). Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow

stromal cell niches. Immunity *25*, 977–988.

[Surh and Sprent, 2008] Surh, C. D. and Sprent, J. (2008). Homeostasis of naive and memory T cells. Immunity *29*, 848–862.

[Taichman et al., 1996] Taichman, R. S., Reilly, M. J. and Emerson, S. G. (1996). Human osteoblasts support human hematopoietic progenitor cells in vitro bone marrow cultures. Blood *87*, 518–524.

[Tussiwand et al., 2011] Tussiwand, R., Engdahl, C., Gehre, N., Bosco, N., Ceredig, R. and Rolink, A. G. (2011). The preTCR-dependent DN3 to DP transition requires Notch signaling, is improved by CXCL12 signaling and is inhibited by IL-7 signaling. Eur. J. Immunol. *41*, 3371–3380.

[Tyson et al., 2012] Tyson, D. R., Garbett, S. P., Frick, P. L. and Quaranta, V. (2012). Fractional proliferation: a method to deconvolve cell population dynamics from single-cell data. Nat. Methods *9*, 923–928.

[Vibert and Thomas-Vaslin, 2017] Vibert, J. and Thomas-Vaslin, V. (2017). Modelling T cell proliferation: Dynamics heterogeneity depending on cell differentiation, age, and genetic background. PLoS Comput. Biol. *13*, e1005417.

[White et al., 1990] White, R. A., Terry, N. H. and Meistrich, M. L. (1990). New methods for calculating kinetic properties of cells in vitro using pulse labelling with bromodeoxyuridine. Cell Tissue Kinet *23*, 561–573.

[Wilson et al., 2008] Wilson, A., Laurenti, E., Oser, G., van der Wath, R. C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C. F., Eshkind, L., Bockamp, E., Li, P., Macdonald, H. R. and Trumpp, A. (2008). Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. Cell *135*, 1118–1129.

[Xing et al., 2004] Xing, Z., Conway, E. M., Kang, C. and Winoto, A. (2004). Essential role of survivin, an inhibitor of apoptosis protein, in T cell development, maturation, and homeostasis. J. Exp. Med. *199*, 69–80.

[Xiong et al., 2013] Xiong, J., Parker, B. L., Dalheimer, S. L. and Yankee, T. M. (2013). Interleukin-7 supports survival of T-cell receptor-Îš-expressing CD4(-) CD8(-) double-negative thymocytes.

Immunology *138*, 382–391.

[Yang et al., 2005] Yang, L., Bryder, D., Adolfsson, J., Nygren, J., M?nsson, R., Sigvardsson, M. and Jacobsen, S. E. (2005). Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. Blood *105*, 2717–2723.

[Yates, 2014] Yates, A. J. (2014). Theories and quantification of thymic selection. Front Immunol *5*, 13.

[Zerjatke et al., 2017] Zerjatke, T., Gak, I. A., Kirova, D., Fuhrmann, M., Daniel, K., Gonciarz, M., Mueller, D., Glauche, I. and Mansfeld, J. (2017). Quantitative Cell Cycle Analysis Based on an Endogenous All-in-One Reporter for Cell Tracking and Classification. Cell Rep *19*, 1953–1966.

[Zhang et al., 2003] Zhang, J., Niu, C., Ye, L., Huang, H., He, X., Tong, W. G., Ross, J., Haug, J., Johnson, T., Feng, J. Q., Harris, S., Wiedemann, L. M., Mishina, Y. and Li, L. (2003). Identification of the haematopoietic stem cell niche and control of the niche size. Nature *425*, 836–841.

[Zhao et al., 2019] Zhao, M., Tao, F., Venkatraman, A., Li, Z., Smith, S. E., Unruh, J., Chen, S., Ward, C., Qian, P., Perry, J. M., Marshall, H., Wang, J., He, X. C. and Li, L. (2019). N-Cadherin-Expressing Bone and Marrow Stromal Progenitor Cells Maintain Reserve Hematopoietic Stem Cells. Cell Rep *26*, 652–669.

[Zhou et al., 2015] Zhou, B. O., Ding, L. and Morrison, S. J. (2015). Hematopoietic stem and progenitor cells regulate the regeneration of their niche by secreting Angiopoietin-1. Elife *4*, e05521.

[Zhou et al., 2014] Zhou, B. O., Yue, R., Murphy, M. M., Peyer, J. G. and Morrison, S. J. (2014). Leptin-receptor-expressing mesenchymal stromal cells represent the main source of bone formed by adult bone marrow. Cell Stem Cell *15*, 154–168.

Acknowledgments

I will of course thank everyone who was directly involved in the pursuit of this thesis but first I would like to look back a little bit.

I started writing these acknowledgments almost 10 years ago.

It was at midday, in late August 2010, at that time I was working as a legal adviser for a French bank near Gare dAusterlitz in Paris. Following a short lunch, I crossed Le Jardin des Plantes, home of Buffon, Cuvier and other prominent French naturalists, a community I have yearned to be part of since then. I reached the office of Claude Gazeau, who was in charge of the evening classes at the biology faculty of Université Paris VI. I was wearing a suit, I introduced myself as a legal adviser in investment banking, declared with honesty that I knew very little about science, having graduated with a high school diploma with major in literature and foreign languages and then having pursued studies in law. Maybe she needed to fill some seats to keep the progamme running, or maybe she recognized I had a real craving for science, anyway she accepted me in the bachelor programme, and I guess from that point on, my path was all set. Thank you Mrs. Gazeau for sharing your passion for plants, and in general thank you for opening my eyes to the incredible diversity of life. I will not (completely) forget about haplodiplontic life cycles of fungi or the reproductive apparatus of monocotyledon flowers. Every day I try to keep in mind that life is wonderful because weird and apparently inefficient.

And now back to the more direct contributions.

My deepest thanks go to Thomas Höfer for giving me the chance to integrate his team of highly spirited scientists. I remember very well my first visit of his lab. It was one of the few instances in my life when I've known without a doubt that I wanted to be part of it. Thank you Thomas for the

stimulating scientific discussions, and the advice during my PhD.

I thank Hans-Reimer Rodewald for accepting me as a guest in his lab, and for the constructive comments and advice on my work.

I thank the members of my thesis advisory committee Ana Martin-Villalba and Lazaro Centanin. Thank you for your encouragements and enthusiasm. Each TAC meeting was a source of inspiration and energy, each time I felt regenerated.

I thank Melania Barile who really got me started with parameter estimation and Nils Becker for his rigorous and meticulous approach to modeling, I have really enjoyed building the cell cycle method with you.

In general, I thank all the colleagues from the Höfer group for the scientific and non scientific discussions. In the office, at the Mensa or in a restaurant, good company is all you need.

I thank every members of the Rodewald lab and in particular Ann-Katrin Fanti and Csilla Kongsaysak-Lengyel. Thank you Anni and Csilla for finding time to help me with my experiments and for listening to my hypotheses and ideas often delivered in my typical overenthusiastic way.

I thank Thomas Höfer and Ana Martin-Villalba for reviewing my thesis and I thank Lazaro Centanin and Michael Milsom for agreeing to be part of my defence committee.

Finally, I thank my parents, brothers and my wife for their moral support. There are a good many sources of stress during a PhD and our conversations helped me to go through the occasional emotional strains.