Lineage-affiliated specification events and molecular characterization during early stages of haematopoietic development

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
Vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Llucia Inès Albertí Servera
aus Palma (Mallorca), Spain

Basel, 2017
Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Antonius Rolink
Prof. Dr. Daniela Finke

Basel, den 15.11.2016

Prof. Dr. Jörg Schibler
Dekan
“Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.”

Marie Skłodowska Curie
# Table of Contents

1. Abbreviations ........................................................................................................... 1

2. Summary ..................................................................................................................... 5

3. Introduction .................................................................................................................. 7
   3.1 Initiation of haematopoiesis ............................................................................... 7
   3.2 Early Progenitor with Lymphoid and Myeloid developmental potential (EPLM) ....... 9
   3.3 Pair-wise model of haematopoiesis ..................................................................... 11
   3.4 Single-cell RNA sequencing: the method that dissects heterogeneity ................. 13
      3.4.1 Technological advances ............................................................................. 14
      3.4.2 Applications ............................................................................................ 15
   3.5 The role of IL-7 and Flt3L cytokines in lymphocyte development ....................... 17
   3.6 Flt3Ltg mouse .................................................................................................... 19

4. Aim of the project ....................................................................................................... 21

5. Materials and methods .............................................................................................. 23
   5.1 Mice .................................................................................................................... 23
   5.2 Flow cytometry and cell sorting ......................................................................... 23
   5.3 In vitro limiting dilution assay .......................................................................... 24
   5.4 In vivo reconstitution assay ............................................................................... 24
   5.5 Bulk cultures with cytokines ............................................................................ 25
   5.6 Statistical analysis ............................................................................................. 25
   5.7 Bulk RNA sequencing ....................................................................................... 25
      5.7.1 Sample acquisition .................................................................................... 25
      5.7.2 RNA extraction ........................................................................................ 25
      5.7.3 RNA quality ............................................................................................. 26
      5.7.4 Library preparation .................................................................................... 27
      5.7.5 Sequencing ............................................................................................... 27
      5.7.6 Pre-processing of sequencing data ............................................................ 28
      5.7.7 Data analysis ............................................................................................ 29
   5.8 Single-cell RNA sequencing .............................................................................. 29
      5.8.1 Capture of single cells ............................................................................... 29
      5.8.2 Library preparation .................................................................................... 30
      5.8.3 Sequencing ............................................................................................... 31
5.8.4 Pre-processing of sequencing data ................................................................. 32
5.8.5 Data analysis ................................................................................................. 32

6. Results .................................................................................................................. 35
6.1 EPLM progenitor population can be divided into at least 4 subpopulations .......... 35
6.2 The heterogeneity of EPLM is reflected in different sets of potentials for each
subpopulation ........................................................................................................... 38
6.3 Ly6D⁺ EPLM has a lymphoid whereas TN a myeloid genetic signature .......... 42
6.4 EPLM subpopulations are developmentally related ......................................... 48
6.5 Single-cell RNA sequencing reveals distinct degree of molecular heterogeneity of
Ly6D⁺ and TN EPLM subpopulations ...................................................................... 49
6.6 Identification of two Ly6D⁺ and three TN subgroups with distinct genetic signatures
....................................................................................................................................... 51
6.7 Expression of lymphoid and myeloid genes is mutually exclusive in single EPLM cells
....................................................................................................................................... 56
6.8 Selective markers for each subgroup. TdT the best candidate to separate cells with
lymphoid from those with myeloid genetic profiles ................................................. 59
6.9 Excluding CD115⁺ cells enriches for TN cells with lymphoid profile ............... 61
6.10 Ebf1 enriches for Ly6D⁺ cells with B-cell potential ......................................... 63

7. Discussion .............................................................................................................. 67

8. References ............................................................................................................ 77

9. Appendix .............................................................................................................. 87
9.1 Paper 1: Permissive roles of cytokines Interleukin-7 and Flt3-ligand in mouse B cell
lineage commitment .................................................................................................. 89
9.2 Paper 2: Reconstitution of a functional B-cell compartment in immunodeficient mice
with pro-B cells propagated with or without stromal cells ...................................... 129
9.3 Review: Versatility of stem and progenitor cells and the instructive actions of
cytokines shape haematopoiesis ............................................................................. 161

10. Acknowledgements ............................................................................................. 175
1. Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACK</td>
<td>ammonium-chloride-potassium</td>
</tr>
<tr>
<td>AML</td>
<td>acute myeloid leukemia</td>
</tr>
<tr>
<td>Batf3</td>
<td>basic leucine zipper transcriptional factor ATF-like 3</td>
</tr>
<tr>
<td>BCR</td>
<td>B-cell receptor</td>
</tr>
<tr>
<td>Blk</td>
<td>B lymphocyte kinase</td>
</tr>
<tr>
<td>BM</td>
<td>bone marrow</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>ccr2</td>
<td>chemokine receptor 2</td>
</tr>
<tr>
<td>CD</td>
<td>cluster of differentiation antigen</td>
</tr>
<tr>
<td>cDC</td>
<td>conventional DC</td>
</tr>
<tr>
<td>Cebpa</td>
<td>CCAAT/enhancer-binding protein alpha</td>
</tr>
<tr>
<td>Ciita</td>
<td>class II transactivator</td>
</tr>
<tr>
<td>CLP</td>
<td>common lymphoid progenitor</td>
</tr>
<tr>
<td>CMP</td>
<td>common myeloid progenitor</td>
</tr>
<tr>
<td>cor</td>
<td>correlation</td>
</tr>
<tr>
<td>CPM</td>
<td>counts per million</td>
</tr>
<tr>
<td>Csf1r</td>
<td>colony-stimulating factor 1 receptor</td>
</tr>
<tr>
<td>Ctla</td>
<td>cytotoxic T-lymphocyte-associate antigen</td>
</tr>
<tr>
<td>Cts</td>
<td>cathepsin</td>
</tr>
<tr>
<td>DC</td>
<td>dendritic cell</td>
</tr>
<tr>
<td>DEG</td>
<td>differentially expressed genes</td>
</tr>
<tr>
<td>DL1</td>
<td>delta-like 1</td>
</tr>
<tr>
<td>DN</td>
<td>double-negative</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>Dnnt</td>
<td>DNA nucleotidyltransferase</td>
</tr>
<tr>
<td>DP</td>
<td>double positive</td>
</tr>
<tr>
<td>dsDNA</td>
<td>double strand DNA</td>
</tr>
<tr>
<td>Ebf1</td>
<td>early B-cell factor 1</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EP</td>
<td>erythrocyte progenitor</td>
</tr>
<tr>
<td>EPLM</td>
<td>early progenitor with lymphoid and myeloid potential</td>
</tr>
<tr>
<td>ERC</td>
<td>extrachromosomal rDNA circles</td>
</tr>
<tr>
<td>F</td>
<td>fraction</td>
</tr>
<tr>
<td>FC</td>
<td>fold change</td>
</tr>
<tr>
<td>FACS</td>
<td>fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FB</td>
<td>follicular B cell</td>
</tr>
<tr>
<td>FcR</td>
<td>fragment crystallizable receptor</td>
</tr>
<tr>
<td>FDR</td>
<td>false discovery rate</td>
</tr>
<tr>
<td>FGF4</td>
<td>fibroblast growth factor-4</td>
</tr>
<tr>
<td>Flt3</td>
<td>Fms-like tyrosine kinase 3</td>
</tr>
<tr>
<td>Flt3L</td>
<td>Flt3 ligand</td>
</tr>
<tr>
<td>Flt3Ltg</td>
<td>human Flt3L transgenic</td>
</tr>
<tr>
<td>FPKM</td>
<td>fragments per kilobase of transcript per million mapped reads</td>
</tr>
<tr>
<td>GO</td>
<td>gene ontology</td>
</tr>
<tr>
<td>HSC</td>
<td>haematopoietic stem cell</td>
</tr>
<tr>
<td>Igα</td>
<td>immunoglobulin alfa</td>
</tr>
<tr>
<td>Igll1</td>
<td>immunoglobulin lambda-like polypeptide 1</td>
</tr>
<tr>
<td>IL</td>
<td>interleukin</td>
</tr>
<tr>
<td>IL7r</td>
<td>IL7 receptor</td>
</tr>
<tr>
<td>IL7Rα</td>
<td>IL7 receptor alfa chain</td>
</tr>
<tr>
<td>IRES</td>
<td>internal ribosomal entry sequence</td>
</tr>
<tr>
<td>Itgb7</td>
<td>integrin beta-7</td>
</tr>
<tr>
<td>Lat</td>
<td>linker for activation of T-cells</td>
</tr>
<tr>
<td>Lax</td>
<td>linker for activation of X cells</td>
</tr>
<tr>
<td>Lck</td>
<td>leukocyte C-terminal Src kinase</td>
</tr>
<tr>
<td>LD</td>
<td>limiting dilution</td>
</tr>
<tr>
<td>Lin</td>
<td>lineage</td>
</tr>
<tr>
<td>LMPP</td>
<td>lymphoid primed multipotent progenitor</td>
</tr>
<tr>
<td>LSK</td>
<td>Lin&lt;sup&gt;-&lt;/sup&gt; Sca-1&lt;sup&gt;+&lt;/sup&gt; c-Kit&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>LT-HSC</td>
<td>long term HSC</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Ly</td>
<td>lymphocyte antigen</td>
</tr>
<tr>
<td>March1</td>
<td>membrane-associated RING-CH protein I</td>
</tr>
<tr>
<td>M-CSF</td>
<td>macrophage colony-stimulating factor</td>
</tr>
<tr>
<td>MEP</td>
<td>megakaryocyte-erythroid progenitor</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>MHCII</td>
<td>MHC class II</td>
</tr>
<tr>
<td>Mpo</td>
<td>myeloperoxidase</td>
</tr>
<tr>
<td>MPP</td>
<td>multipotent progenitor</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>MZB</td>
<td>marginal zone B cell</td>
</tr>
<tr>
<td>NGS</td>
<td>next generation sequencing</td>
</tr>
<tr>
<td>NK</td>
<td>natural killer</td>
</tr>
<tr>
<td>NSG</td>
<td>NOD scid gamma</td>
</tr>
<tr>
<td>PAM</td>
<td>partioning around medoids</td>
</tr>
<tr>
<td>Pax5</td>
<td>paired box protein 5</td>
</tr>
<tr>
<td>PBMC</td>
<td>peripheral blood mononuclear cells</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate-buffered saline</td>
</tr>
<tr>
<td>PCA</td>
<td>principal component analysis</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>pDC</td>
<td>plasmacytoid DC</td>
</tr>
<tr>
<td>Prtn3</td>
<td>proteinaise 3</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>quantitative reverse transcription PCR</td>
</tr>
<tr>
<td>QS</td>
<td>quality score</td>
</tr>
<tr>
<td>Rag</td>
<td>recombination-activating gene</td>
</tr>
<tr>
<td>RIN</td>
<td>RNA integrity number</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNA-seq</td>
<td>RNA sequencing</td>
</tr>
<tr>
<td>SCF</td>
<td>stem-cell factor</td>
</tr>
<tr>
<td>sd</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SDF1</td>
<td>stromal cell-derived factor-1</td>
</tr>
<tr>
<td>SEM</td>
<td>standard error of the mean</td>
</tr>
<tr>
<td>SiglecH</td>
<td>sialic acid binding Ig-like lectin H</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Sla2</td>
<td>src-like-adaptor protein</td>
</tr>
<tr>
<td>Stat5</td>
<td>signal transducer and activator of transcription 5</td>
</tr>
<tr>
<td>ST-HSC</td>
<td>short term HSC</td>
</tr>
<tr>
<td>TCR</td>
<td>T-cell receptor</td>
</tr>
<tr>
<td>TdT</td>
<td>terminal deoxynucleotide transferase</td>
</tr>
<tr>
<td>TF</td>
<td>transcription factor</td>
</tr>
<tr>
<td>Th2</td>
<td>T helper cell type 2</td>
</tr>
<tr>
<td>Tlr</td>
<td>toll-like receptor</td>
</tr>
<tr>
<td>TN</td>
<td>triple negative</td>
</tr>
<tr>
<td>Trat1</td>
<td>T-cell receptor-associated transmembrane adapter 1</td>
</tr>
<tr>
<td>UMI</td>
<td>unique molecular identifiers</td>
</tr>
<tr>
<td>VCAM-1</td>
<td>vascular cell-adhesion molecule-1</td>
</tr>
<tr>
<td>vs</td>
<td>versus</td>
</tr>
<tr>
<td>WT</td>
<td>wild type</td>
</tr>
<tr>
<td>Zap70</td>
<td>70 kDa zeta-chain associated protein</td>
</tr>
</tbody>
</table>
2. Summary

The well-established “classical” model of haematopoiesis reflects a hierarchical decision-making process where early multipotent progenitors make an irrevocable decision to differentiate towards either the lymphoid or myeloid lineages through so-called Common Lymphoid and Common Myeloid progenitor intermediates respectively. However, the proposals for alternative differentiation pathways and the description of progenitor cells that contradict the lympho-myeloid dichotomy have prompted multiple revisions of the strict compartmentalized classical model. We have previously characterized a B220⁺ CD117⁺ CD19⁻ NK1.1⁻ uncommitted and multipotent haematopoietic progenitor with combined lymphoid and myeloid potential that we called Early Progenitor with Lymphoid and Myeloid potential (EPLM). The emergence of high throughput methods enabling the investigation of single-cell whole-transcriptome profiles generates data that enhances the active debate regarding the heterogeneity of apparently phenotypically homogenous progenitors having different multiple lineage potentials. This thesis provides a detailed analysis of EPLM heterogeneity by combining the alternative and complementary “top-down” and “bottom-up” experimental designs. Using the “top-down” approach based on the expression of the cell-surface markers Ly6D, SiglecH and CD11c, we could subdivide EPLM into four subpopulations with distinct lineage biases. As revealed by the subsequent functional experiments, the Ly6D⁺ EPLM fraction was lymphoid restricted and contained most B-cell potential whereas the so-called triple negative (TN) EPLM expressing none of the above markers remained as a lympho-myeloid multipotent fraction and the potential precursor of the Ly6D⁺ subset. Subsequently, single-cell RNA sequencing (“bottom-up” approach) of 152 Ly6D⁺ and 213 TN single cells revealed that in fact TN are composed of a mixture of cells where the myeloid potential is mainly due to the contribution of the G3 TN subset whereas the lymphoid potential resides in the G1 TN clustered group of cells. This heterogeneity was masked in previous bulk molecular and functional experiments, thus demonstrating the power of single-cell RNA-sequencing technology to study heterogeneity in haematopoietic progenitors at an unprecedented resolution. Moreover, single-cell transcriptome profiles enabled
the detection in an unbiased manner of markers that better define cellular identity. Here we redefined the “top down” EPLM classification by identifying Terminal deoxynucleotide Transferase (TdT) as a potential marker with which to discriminate the lymphoid and myeloid potential of EPLM since, in addition to the previously identified lymphoid primed Ly6D$^+$ cells, TdT is also expressed in the G1 TN fraction, which turned to be molecularly indistinguishable from the G2 Ly6D$^+$ fraction. The use of other candidate markers such as Ebf1 and CD115 enabled us to prospectively isolate cells from different newly identified subgroups of EPLM and to confirm their genetic signatures with functional assays, thus supporting the increasing belief that the repertoire of genes expressed reflects the immediate lineage bias of that cell. Finally, within the Ly6D$^+$ cells, we found a B-cell priming gradient and propose that the G1 Ly6D$^+$ fraction is the direct precursor of the first B-cell committed stage, the CD117$^+$ CD19$^+$ Pro-B cell. Therefore, we favour the concept that haematopoiesis occurs through a process of graded commitment where molecular priming is initiated earlier than previously anticipated. Overall, this study provides a valuable model demonstrating that previously characterized, phenotypically homogeneous, multipotent progenitor cells are in fact composed of mixtures of cells with differently restricted differentiation capacities.
3. Introduction

3.1 Initiation of haematopoiesis

Haematopoiesis is a well-orchestrated and tightly regulated process resulting in the formation of all blood cell types. In adult mammals, haematopoiesis starts in the bone marrow (BM) with a rare cell type called Haematopoietic Stem Cell (HSC) that has both self-renewal and multipotent capacities. Self-renewal is the ability to generate cells that are identical to the mother HSCs without differentiation, thus allowing the maintenance of the stem cell pool size, while multipotency is the ability to differentiate into all functional blood cells.

Self-renewal and differentiation of HSC requires a special environment, termed the hematopoietic stem cell niche, which provides the protection and the necessary signals for their maintenance and development [1, 2]. Two main microenvironments are suggested to form the bone marrow niches: the endosteal and the vascular niche [3]. The endosteal niche, located in the interface between bone and BM [4], includes osteoblasts, derived from mesenchymal precursors, that produce a number of cell-signalling molecules that contribute to the maintenance and regulation of HSC such as Jagged (ligand of Notch receptor), thrombopoietin and angiopoietin, SCF, chemokines or Wnt [5]. The vascular niche, located at the fenestrated endothelium of bone marrow sinusoids [4, 6], promotes proliferation and differentiation, active cycling and generation of short-term HSCs by producing factors important for mobilization, homing and engraftment of HSC such as vascular cell-adhesion molecule-1 (VCAM-1), stromal cell-derived factor-1 (SDF1) or fibroblast growth factor-4 (FGF4) [7]. An important function of the stem cell niche is to regulate the balance between cellular self-renewal and differentiation [2]. One mechanism that ensures this balance is the control of asymmetric/symmetric cell division. As a result of asymmetric division, one daughter cell maintains the mother stem cell fate and the other daughter cell becomes more committed to terminal differentiation [8, 9]. Several reported mechanisms argue for an intrinsic capability of HSCs to divide asymmetrically, exemplified by the unequal molecular segregation during cell division. Some examples are: the asymmetric distribution of regulatory
cell surface molecules (e.g. notch [10]) or cellular processes [11] in mammalian neurogenesis; asymmetric distribution of cellular organelles (e.g. p-granules in C.elegans [12]); asymmetric distribution of transcription factors and co-factors (e.g. Prospero/Miranda in Drosophila [13]); and asymmetric distribution of DNA (e.g. extrachromosomal rDNA circles, ERCs, in yeast [14]). Alternative to the intrinsic model, there is a post-mitotic model arguing that the observed differences in the cell fates of sister cells are due to post-mitotic mechanisms in which initially identically specified daughter cells influence each other’s cell fate (e.g. via Notch signalling pathway [15, 16]). Further investigations are required to resolve whether both models co-exist and to what extent each model contributes to different specification of daughter cells.

Multi-potentiality of HSCs was first demonstrated by Osawa et al. after showing that a single mouse CD34<sup>low</sup> HSC reconstitutes haematopoiesis long-term in a lethally irradiated mouse [17, 18]. Similar results were reported by Notta et al. with single human HSCs engrafting NGS mouse [19]. Since most mature blood cell types are short lived, HSCs are ultimately responsible for the continuous and life-long renewal of blood. Mature blood cells are produced at an impressive rate of more than one million cells per second in the adult human.

The description of HSCs, and their isolation by fluorescence activated cell sorting (FACS), relies on the presence or absence of a range of cell-surface molecules. HSCs do not express markers that are associated with the various haematopoietic cell lineages (Lin<sup>-</sup>) including, for example, CD3 (T lymphocytes), B220 and CD19 (B lymphocytes), CD11b (monocytes-macrophages), Ly-6G (neutrophils) and TER-119 (erythroid cells), and they express the two molecules c-Kit, a tyrosine kinase receptor for the cytokine stem cell factor (SCF), and Sca-1 a phosphatidinositol-anchored membrane protein of unknown function [20]. Thus, HSCs are enriched in the Lin<sup>-</sup>, Sca-1<sup>-</sup>, c-Kit<sup>+</sup> fraction of bone marrow cells, termed LSK cells, that represent approximately 0.05% of nucleated adult mouse BM cells [21]. The LSK compartment contains cells that repopulate the entire haematopoietic system either long-term (LT-HSC) or short-term (ST-HSC). Three different FACS staining strategies describe the phenotype of the two LSK subsets. The original Weissman approach relies on Thy-1.1 and Flt3 (fms-like) tyrosine kinase, a type 3-
receptor kinase, whose expression was observed to correlate with a loss of the capacity of HSCs to self-renew [22, 23]. In this model LT-HSCs are LSK Flt3\(^-\) Thy1\(^{low}\) and, ST-HSCs are LSK Flt3\(^+\) Thy1\(^{low}\). Later, in 2005, Jacobsen combined LSK markers with CD34 and Flt3 [22], a commonly used strategy. Thus Jacobsen et al. described LT-HSCs as LSK CD34\(^-\) Flt3\(^-\) and ST-HSC as LSK CD34\(^+\) Flt3\(^-\). Recently, Morrison made use of additional signalling lymphocyte activation molecule markers (CD150, CD48, CD229, and CD244) to subdivide LSK cells. Here, LT-HSCs are defined as LSK CD150\(^-\), CD48\(^{-/low}\), CD229\(^{-/low}\) and CD244\(^-\) while ST-HSC differ in CD229 expression and are defined as CD150\(^+\), CD48\(^{-/low}\), CD229\(^+\) and CD244\(^-\) [23].

HSCs differentiate into multipotent progenitor cells (MPPs), which gradually lose their multipotentiality while undergoing decision-making, expansion, and differentiation via intermediate lineage precursors giving rise to the final compartments of functional cells. The long-standing classical model of haematopoiesis [24], together with other versions [25], reflects a hierarchical decision-making process where early multipotent progenitors make an irrevocable decision to differentiate towards either the lymphoid or the myeloid lineage [26].

This clear lympho-myeloid separation is supported by the identification of progenitors of each lineage. Weissman and colleagues reported the so-called common lymphoid progenitor (CLP) [27] which can give rise to B and T lymphocytes and natural killer (NK) cells, and the common myeloid progenitor (CMP) [28] which can differentiate into the rest of the immune and blood cells. Thus, this traditional model illustrates single routes of differentiation to each end-cell type. However, over the past 15 years, the above strict compartmentalization of hematopoietic cells has been challenged. This is mainly due to the description of progenitor cells that contradict the lympho-myeloid dichotomy.

### 3.2 Early Progenitor with Lymphoid and Myeloid developmental potential (EPLM)

The first indication that haematopoietic development has much more plasticity than previously anticipated came from studies performed in mice deficient
for the transcription factor *Pax5* [29]. The *Pax5*−/− mouse presents an absolute block in B-cell development at the Pro-B cell stage. Remarkably, and in contrast to wild-type (WT) cells, *Pax-5* deficient Pro-B cells had multilineage developmental potential since they could develop into myeloid, NK, and T cells both *in vitro* and *in vivo* [30-32]. Moreover, even WT committed precursor B cells regain multipotentiality upon conditional inactivation of the *Pax5* gene [32]. This demonstrates that progenitor cells that have progressed some way along a given lineage retain the ability to give rise to other cell types until a rather advanced stage of differentiation. However, whether such a degree of developmental plasticity occurs under physiological conditions is an unresolved issue. Subsequently, our laboratory detected in wild-type mouse BM a cell with similar properties to the *Pax5*−/− Pro-B cells. This cell has the capacity to differentiate into lymphoid and myeloid cells and was therefore called EPLM (Early Progenitor with Lymphoid and Myeloid developmental potential). EPLM cells were identified as B220+ cKitint CD19− and NK1.1− cells representing 0.2% of all nucleated BM cells [33] (Figure 1). In terms of phenotype, this progenitor is closely related to the CLP with the marked difference of B220 expression, EPLM being B220+ whereas CLP are B220−, and partially overlaps with the so-called Fraction A cells identified by Hardy and co-workers [34] (Table 1 contains a comparison of the nomenclatures used for early B-cell developmental stages). Limiting dilution analysis of EPLMs cultured together with stromal cells and addition of appropriate cytokines, enabled the quantification of *in vitro* B, T and, myeloid precursor frequencies. EPLMs showed strong B-cell developmental potential and strong-to-moderate differentiation potential for T cells and myeloid cells (mostly macrophages).
Therefore, this suggested that under physiological conditions the developmental fate of EPLM is mainly to become B cells. Reconstitution assays in order to assess the EPLM’s in vivo developmental potentials revealed their ability to transiently reconstitute both B and T cell compartments in sublethally irradiated Rag2-deficient mice.

Another cell that contradicts the classical lympho-myeloid lineage separation is the Lymphoid-Primed MultiPotent progenitor (LMPP), which has little potential for megakaryocyte or erythroid development while retaining other potentials [35]. Compared with EPLM, LMPP present slower kinetics in developing into lymphoid cells, hence suggesting an earlier and upstream location in haematopoietic development. Moreover, LMPP is the Flt3 positive high fraction of LSK and, in contrast with EPLM, is B220+. Together with the identification of new progenitor cells, the finding of alternative routes to end-cell types [36] prompted the necessity of proposing new models for the architecture of haematopoiesis.

Table 1. B-cell developmental stages in the bone marrow. Comparison of the nomenclatures used to identify developmental B cell subsets and how they relate to key VDJ recombination events (comprehensively reviewed in Osmond et al. [37], and Hardy et al. [38]). Table adapted from Hardy et al. [39].

<table>
<thead>
<tr>
<th>Developmental stages</th>
<th>Osmond</th>
<th>Melchers and Rolink</th>
<th>Hardy</th>
<th>Status of Ig loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro-B</td>
<td>Pre-pro B / EPLM</td>
<td>A</td>
<td>Germline</td>
<td></td>
</tr>
<tr>
<td>Pro-B</td>
<td>B</td>
<td>D-Jκ rearrangement</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>Vκ-DJκ rearrangement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-B</td>
<td>Large pre B</td>
<td>C’</td>
<td>VκDJκ pairs with λ5-Vpre-B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pre-BCR surface expression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small pre B</td>
<td>D</td>
<td>Vκ-Jκ or Vκ-Jκ rearrangement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immature B</td>
<td>E</td>
<td>Complete BCR (receptor editing can occur)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3 Pair-wise model of haematopoiesis

In collaboration with others, a few years ago our laboratory presented a new model of haematopoiesis called the pair-wise model [40, 41]. In contrast to the traditional hierarchical trees, this model outlines haematopoietic development in a cyclical representation. Fundamentally, the principle of an HSC giving rise to all blood cell types is the same. However, the intermediate progenitor stages are better illustrated. In the pair-wise model, the set of developmental potentials that have
been shown experimentally are represented as a continuum by arcs (Figure 2). As development proceeds, the arcs become shorter indicating that developmental potentials are gradually restricted and ultimately cells become committed to a differentiated cell fate (as shown by the outermost arc). As its name indicates, this model depicts a series of pair-wise developmental relationships between the various haematopoietic lineages. This infers that only certain bi-potentialities are permissive. For example, a bi-potent T lymphocyte-megakaryocyte cell should not exist, and has not been described to date.

Figure 2. Pair-wise model of haematopoiesis. Illustrated is a fate choice continuum with an invariant series of pair-wise developmental relationships between haematopoietic cells’ fates.

The pair-wise model has no lineage branching patterns and therefore, does not assume a single preferred route to a particular end-cell type. In other words, HSCs can reach a specific differentiated blood cell through more than one type of intermediate progenitor. For instance, dendritic cells (DC) can be derived from either CLP/EPLM or CMP, indicating that DC can have both lymphoid and myeloid origin [36]. Moreover, multiple routes have been demonstrated towards neutrophils and monocytes [42, 43]. Another advantage is that the pair-wise model maintains the possibility that progenitors can move in the opposite direction and regain multipotentiality -as is the case for the Pax5−/− Pro-B cells. Overall, the above-described model, which includes the recently identified progenitors and allows
alternative routes, leads to a viewpoint that haematopoiesis is a more versatile and less strictly compartmentalized process than previously thought.

In line with the description of progenitor cells that have multiple lineage options, there is an increasing debate regarding their heterogeneity. It is highly likely that multipotent progenitor cells presently viewed as a homogeneous population will be divided into subsets with lineage biases or even into already committed subsets. In turn, some end-cell types will be probably further divided into numerous subtypes. To what extent progenitor cell populations, and their mature progeny, are heterogeneous is very much dependent on the quantity of new and existing cell-surface markers that can be used to define new subpopulations. In this regard, together with the improvement of traditional flow cytometer-based technologies, allowing the simultaneous detection of an increasing number of markers, the explosion of high throughput technologies, such as RNA sequencing that enable the parallel screening of a large amount of markers in an unbiased manner, has been of a pivotal importance.

3.4 Single-cell RNA sequencing: the method that dissects heterogeneity

In the last couple of years, RNA sequencing at the single-cell level has emerged as an important technology for interrogating the degree of heterogeneity in a population. This technology, which can be performed at a whole-transcriptome scale with unprecedented resolution, consists of extracting the RNA from individual cells followed by a sequencing process similar to that for bulk RNA sequencing. When RNA sequencing is performed at the population level, as has been the case for most studies, we obtain an average transcription profile of thousands or millions pooled cells. This, masks the true distribution of gene expression levels across individual cells, thus hampering the elucidation of whether a subset of cells is responsible for the expression of a given gene or even if it is the major contributor of a function in a seemingly homogeneous population [44]. Therefore, it is crucial to go to the single-cell level where there is no averaging effect.
3.4.1 Technological advances

The first mRNA sequencing whole-transcriptome analysis of a single cell was reported in 2009 by Tang and co-workers. Using a novel mRNA sequencing assay they detected expression of 5,270 genes in a single mouse blastomere [45], thus greatly overcoming the limited throughput of a couple of hundred genes that can be analysed by single-cell qRT-PCR. Moreover, in contrast to qPCR, mRNA sequencing based technology provides a comprehensive view of the transcriptome profile in one cell in an unbiased manner or, in other words, without requiring prior knowledge and selection of genes of interest. However, the method reported relies heavily on manual manipulation of cells and reagents and can only process few cells per hour. The limitation of manual handling, either during the capturing or during the later processing steps, also occurs in other single-cell RNA sequencing technologies such as laser-capture microdissection or single-cell sorting into a multi-well plate. Therefore, researchers rapidly attempted the development of automated methods.

Automation of the single-cell RNA sequencing process became a reality with the emergence of chip-based microfluidics devices. In 2013, Fluidigm launched the “C1™ Single-Cell Auto Prep System” helping expand access to the single-cell RNA sequencing technology [46, 47]. The C1 system provides a completely automated workflow that integrates both the capture of single cells in a microfluidic chip and the later processing steps including washing, lysis, reverse transcription and, PCR pre-amplification. In addition, the major strength of microfluidics is the manipulation of minute volumes, working in a nanoliter reaction scale. As a result, albeit allowing the processing of only a few tens of cells per experiment, this system presents higher reproducibility, more cost-effective reactions and increased sensitivity compared with classical methods [48]. In parallel with its commercialization, the journal “Nature Methods” declared single-cell sequencing as the method of the year 2013. As a consequence, the past few years have seen a burst of papers from laboratories around the world that have developed expertise in single-cell approaches. These papers include new single-cell technologies and provide evidence for its valuable application in many areas.
Regarding the development of new technologies, the state-of-the-art has shifted towards scaling up the number of cells that can be analysed in parallel and in an affordable manner in terms of time and costs. Recently, Klein et al. and Macosko et al. have developed two advanced droplet-based microfluidics methods termed Drop-seq [49] and inDrop sequencing [50] respectively. Both methods, published in the same issue of the journal Cell, massively increase the throughput to several thousand of cells while reducing the cost of sequencing. Moreover, despite there being some differences between the experimental protocols, both publications rely on the same fundamental principle. The cells are separated into nanoliter-sized aqueous droplets in carrier oil as reaction chamber. Each reaction chamber contains microparticles coated with unique cell barcodes and Unique Molecular Identifiers (UMIs). Although droplet microfluidics is not a new technology, this efficient barcoding strategy makes it an attractive method and, altogether, has big potential for commercialization.

To date, apart from droplet-based microfluidics, a variety of other microfluidics techniques for single-cell trapping exist. Some examples are hydrodynamic trap [51], magnetic trap [52, 53], acoustic trap [54], dielectrophoretic trap [55], and optical trap [56]. All this together exemplifies increasing interest for single-cell handling technologies and creates anticipation for exciting biological advances for the coming years.

3.4.2 Applications

Single-cell transcriptomics offers a wide range of applications that have already enabled some advances in a broad variety of fields apart from immunology. Examples include: oncology, with the dissection of tumour cell heterogeneity [57], the identification of subgroups associated with anti-cancer drug resistance [58], and the analyses of rare circulating tumour cells [59] or cancer stem cells; embryology, with new insights into the gene regulatory networks controlling early stages of embryonic development [45, 60-62]; neurobiology, with the unravelling of the complex cellular composition of the mouse hippocampus while uncovering novel markers and cell types [47]; or microbiology, with the opportunity of sequencing
uncultivated microbial eukaryotes at the single-cell level and identifying new species, unknown microbial functions and revised evolutionary relationships between eukaryotes, bacteria and archea [63, 64].

In immunology, single-cell molecular profiling is providing valuable information on the differentiation and function of the immune cells both at the steady state and during immune responses. Moreover, it is opening up a new avenue for finding markers that better describe cell phenotypes in an unbiased manner. During the forthcoming years, it is predicted that the identity of many immune cells will be redefined while possibly new cell types and subtypes will emerge. Indeed, several studies have already pointed in this direction. Mahata et al. identified a Th2 subpopulation contributing to the maintenance of T-cell homeostasis. It is of interest that the Th2 subset, characterized by the expression of a specific enzyme (Cyp11a1), could be isolated by using surface molecules whose expression correlates with enzyme expression, thus allowing for the first time a broad ex vivo functional validation of a new cell type discovered by single-cell RNA sequencing [65]. Similarly, Shalek et al. found a very small subset of cells among mouse bone-marrow-derived dendritic cells. Those cells, termed precocious, are the first ones that produce and secrete a wave of interferon during antiviral responses [66]. Another example with novel subsets is a recently published paper providing a detailed study about the inter-cellular transcriptomic variation within the traditionally classified classical, intermediate and non-classical monocytes [67]. In the publication, Gren et al. reveal that each group of monocytes contains further subgroups with distinct genetic signatures according to their activation status and differentiation. This indicates the ability of single-cell transcriptomics to discover cell heterogeneity within defined cell populations.

However, as a proof of concept, massive parallel single-cell RNA sequencing has also been applied to more complex and heterogeneous systems such as the spleen [68] or peripheral blood mononuclear cells (PBMCs) [53]. Apart from the identification of previously unknown hidden subpopulations, this approach could reconstitute the global cell heterogeneity within splenic tissue and PBMCs, thus envisaging the opportunity to redefine our knowledge on lineage hierarchy. In fact, a number of studies have amplified the current hematopoietic classifications and
differentiation trajectories through the detection of transient or intermediate cell states [69-71]. For instance, Paul et al. identified seven subgroups of c-Kit^+ Sca1^− Lin^− sorted cells representing a gradient of erythrocyte transcription, from expression of early progenitor genes towards the up-regulation of functional genes [69]. It is of note that, unlike current models, none of the seven subsets co-expressed megakaryocyte genes. Therefore, they suggested that the standard gating and sorting for megakaryocyte-erythrocyte progenitors (MEP) might be termed erythrocyte progenitors (EP) instead. This exemplifies the necessity of revising the identity of the different immune cells as well as their subtypes, lineages and composition within tissues.

In conclusion, the studies reported to date have convincingly demonstrated that single-cell RNA sequencing is a powerful approach that can be used in many applications such as exploring cellular heterogeneity in a population, studying differentiation dynamics, redefining cell identity, or identifying rare cell types among others. Excitingly, it is expected that application of this technology will continue to expand as the techniques and analysis tools evolve.

3.5 The role of IL-7 and Flt3L cytokines in lymphocyte development

Cytokines are the pivotal external factors that deliver environmental signals to control haematopoietic cell development. Their mode of action can be viewed as either instructive or permissive [72]. In the instructive model, cytokines act directly on progenitor cells to drive them towards a specific lineage by actively inducing the initiation of a lineage-specific gene program. In contrast, in the permissive model commitment occurs in a cell-autonomous and stochastic manner and cytokines act to selectively allow cells committed to a particular lineage to survive and/or proliferate [73, 74]. The precise instructive versus permissive role of cytokines in haematopoiesis remains controversial [75-77]. Two important cytokines for T- and B-lymphocyte development are IL-7 and Flt3L since their receptors are co-expressed in a narrow window immediately before commitment and mutant mice show impaired lymphoid development [78-81].
IL-7 was originally identified as a growth factor for B-cell progenitors [82]. Later on, IL-7 activity on thymocytes and T-cell survival was also reported [83]. Subsequently, a specific receptor for IL-7, expressed in both B- and T-cell progenitors, was identified [84]. Mice lacking either the cytokine [81] or its receptor (IL-7Rα) [80] have a leaky arrest of T-cell development at the DN2 stage [80, 81] and absence of γδT cells [85], whereas a dramatic defect in B-cell generation with a block at the Pro-B cells stage occurs. Reports with over-expression of the pro-survival gene Bcl2 demonstrated that whereas the T-cell defect is rescued, thus suggesting a permissive role of IL-7 in early T-cell development [86, 87], B-cell development is not re-established [88, 89], arguing for an instructive role of this cytokine in B-cell commitment at least in mice. Interestingly, humans with defects in IL-7 signalling display a normal B-cell development [90, 91]. However, an in vitro system has provided clear evidence that IL-7R signalling through Stat5 alone is not sufficient to induce transcription of Ebf1 and Pax5 [92], indispensable for B-cell commitment, therefore arguing for an important but redundant action of IL-7 during commitment to the B-cell lineage [93, 94]. Therefore, while the importance of IL-7 as a survival and growth factor for committed B-cell progenitors is well established, its instructive role through Ebf1 and Pax5 up-regulation remains unclear.

Flt3 ligand (Flt3L) is one of the essential cytokines for cell survival and proliferation during early haematopoiesis [95, 96]. Its function has gained much attention since mutations in Flt3L signalling are commonly found in Acute Myeloid Leukaemia (AML) [97]. Using flow cytometry, its receptor (Flt3) has been detected as early as the non-self-renewing MPP stage of development [98]. Thereafter, several downstream myeloid and lymphoid progenitors continue to express Flt3 with the exception of megakaryocyte-erythocyte progenitors [99]. Flt3 expression is extinguished upon lineage commitment and, among differentiated cells, only dendritic cells retain Flt3 on their surface. For instance, during B-cell commitment Pax5 expression directly antagonizes that of Flt3 [100]. In the thymus, Flt3 expression is used to identify the earliest and uncommitted thymocyte subpopulation with remaining B-cell developmental potential [101-103]. Mice with targeted gene disruption of Flt3 [78] or its ligand [79] do not exhibit a complete loss of any haematopoietic population but reduced numbers of B, DC and NK cells [78,
Moreover, upon transplantation, BM progenitors from Flt3$^{-/-}$ or Flt3L$^{-/-}$ mice reconstitute the B-cell compartment poorly [78, 79] and Flt3L was found to be essential for maintaining normal numbers of uncommitted B-cell progenitors [104]. Therefore, these results suggest a survival and proliferative action of Flt3L during B-cell development, rather than instructive. It is probable that Flt3L signalling exerts its function in concert with signalling from other cytokines such as SCF or IL-7 [92, 105]. This is manifested in cultures of lymphoid progenitor cells, such as ETP, CLP or EPLM, in the presence of IL-7, where Flt3L provides and additive anti-apoptotic effect while stimulating proliferation [106, 107].

### 3.6 Flt3Ltg mouse

Until recently, both bulk and single-cell RNA sequencing required a considerable amount of starting material. For bulk experiments, it was about 500 nanograms of RNA per sample, while for the single-cell capturing step using the C1 platform the starting density recommended in the “single-cell preparation guide” from Fluidigm is 166-255 cells/$\mu$L and 5-20 $\mu$L are loaded into the chip. To overcome this limitation, because the EPLM is a rare population, we made use of a mouse model, the Flt3 ligand transgenic (Flt3Ltg) mouse, to isolate EPLM and other progenitor populations in greater numbers.

Initially, Ceredig et al. showed that apart from DC, increased in vivo Flt3L availability after several injections led to transient expansion of Flt3$^+$ progenitors with either myeloid or lymphoid developmental potential or both [108]. Later on, our laboratory generated a transgenic mouse with sustained over-expression of human Flt3L, thus providing for the first time a detailed in vivo analysis of the effect of this cytokine on different haematopoietic lineages [109]. Examination of Flt3Ltg mice confirmed the importance of Flt3L in DC development. Moreover, the transgenic mice presented a tremendous expansion of almost all haematopoietic progenitors in the BM with the exception of MEP. The decrease in MEP occurred in a rapid fashion when WT mice were injected with recombinant Flt3L, therefore suggesting for the first time an instructive role for the Flt3L cytokine. Tsapogas and co-workers proposed that an increased level of Flt3L guides haematopoietic
development towards the lymphoid-myeloid fate at the expense of the megakaryocyte-erythrocyte fate [41, 109]. This provides an explanation for the decreased platelet counts and the consequent development of anaemia in Flt3Ltg mice.

The EPLM compartment is a good example of the tremendous expansion of progenitors in the BM of Flt3Ltg mice. In fact, the number of EPLM cells was increased by 14-fold when compared to WT mice (Figure 3). Therefore, the Flt3Ltg mouse is an excellent tool for the isolation of large numbers of EPLM and other progenitor cells for further in vitro and in vivo experiments, transcriptome profiling or molecular biology analyses.

Figure 3. Total numbers of CMP, CLP and EPLM in the BM of WT (white symbols) and Flt3Ltg (black symbols) mice (5-7 mice per group). ***P<0.0001.
4. Aim of the project

In this thesis we sought to characterize a previously identified uncommitted and multipotent progenitor cell called EPLM with the aim to:

- Dissect EPLM heterogeneity by combining the use of cell surface markers, such as Ly6D, SiglecH and CD11c, the so-called “top down” experimental strategy, with single-cell RNA sequencing, “bottom up” approach.
- Unravel whether EPLM phenotypic or molecular heterogeneity causes different sets of potentials in the identified subpopulations by in vitro limiting dilution assays and in vivo reconstitution experiments.
- Study the precursor-product relationship among the subpopulations in bulk culture with cytokines.
- Identify markers that better define cellular identity of the EPLM subpopulations in an unbiased manner, by single-cell RNA sequencing.
- Utilize the identification of EPLM subpopulations in order to study the role of cytokines (Flt3L and IL-7) in B-cell commitment (von Muenchow et al. Appendix paper 1).
5. Materials and methods

5.1 Mice

C57BL/6 (B6), B6 Rag2-deficient [110], B6 FLT3L transgenic (Flt3Ltg, [109]) and Ebf1<sub>hCd2/+</sub>-Flt3Ltg mice used herein were 6 to 11 weeks old and matched by age and sex for each experiment. All mice were bred and maintained in our animal facility under specific pathogen-free conditions. All animal experiments were carried out according to institutional guidelines (authorization numbers 1886 and 1888 from Kantonales Veterinäramt, Basel).

Ebf1<sub>hCd2/+</sub> mice were provided by Prof. Meinrad Busslinger and co-workers [111]. These mice had been generated by inserting an internal ribosome entry sequence (IRES)-hCd2 (ihCd2) reporter gene into the 3’ untranslated region of the Ebf1 gene. We next crossed the Ebf1<sub>hCd2/+</sub> reporter to the Flt3Ltg mice, thereby generating the Ebf1<sub>hCd2/+</sub>-Flt3Ltg mice.

5.2 Flow cytometry and cell sorting

Bone marrow cell suspensions were obtained from femurs of the two hind legs of mice. Bones were flushed with a 2ml syringe filled with PBS containing 0.5% BSA and 5mM EDTA. Afterwards, single-cell suspensions were subjected to ACK treatment for erythrocyte depletion, stained with the appropriate combination of antibodies for 30 minutes at 4°C, and washed for subsequent flow cytometry or cell sorting. The following antibodies were used (from BD Pharmingen, 18 eBioscience, BioLegend, or produced in house) with names in brackets describing the corresponding clone: anti-B220 (RA3-6B2), anti-CD117 (c2B8), anti-CD19 (1D3), anti-NK1.1 (PK136), anti-SiglecH (551), anti-CD11c (HL3), anti-Ly6D (49-H4), anti-CD115 (AFS98), anti-hCD2 (RPA-2.10) conjugated with FITC, PE, PE/Cy7, APC, Bv421 or Biotin plus streptavidin-Bv650. For detection of TdT, cells were fixed and permeabilized after cell-surface staining using the Foxp3 Fix/Perm buffer kit (eBioscience), and subsequently stained with APC-conjugated anti-TdT (19-3) according to the supplier’s protocol. Flow cytometry was performed using a BD LSR
Fortessa (BD Biosciences) and data were analyzed using FlowJo v9.8 Software (Treestar). For cell sorting, a FACS Aria IIu (BD Biosciences) was used (>98% purity).

5.3 **In vitro** limiting dilution assay

ST2 [112], OP9 [113] and OP9 stromal cells expressing the Notch ligand Delta-like 1 (OP9-DL1) [114] were plated in a 96-well flat-bottom plate one day prior to coculture at 3x10^3 cells per well. The following day, semi-confluent stromal cells were γ-irradiated with 3000 rad using a Cobalt source (Gammacell 40, Atomic Energy of Canada, Ltd) at 100 rad/min and co-cultured with graded numbers of sorted haematopoietic progenitors in 48 replicates (or as indicated). Cells were maintained as a monolayer in IMDM supplemented with 5x10^{-5}M β-mercaptoethanol, 1mM glutamine, 0.03% w/v Primatone (Quest Naarden, The Netherlands), 100U/mL Penicillin, 100 μg/mL Streptomycin and 5% FBS (Amimed) at 37°C in a humidified atmosphere containing 10% CO₂ in the air. OP9 and OP9-DL1 co-cultures were additionally supplemented with 10% IL7-conditioned medium. After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1 and ST2 cell cultures), wells were inspected using an inverted microscope. Wells containing colonies of more than 50 cells were scored as positive. For each experiment, the frequency of negative wells was plotted against the number of haematopoietic progenitors plated and the fraction of progenitor cells developing B-cell, T-cell or myeloid colonies was estimated considering plating efficiency that follows a Poisson distribution.

5.4 **In vivo** reconstitution assay

Recipient Rag2-deficient mice were γ-irradiated using a Cobalt source as previously described at a dose of 400 rad 4h prior to reconstitution. Indicated numbers of sorted haematopoietic progenitors from donor mice (B6, Flt3Ltg or Ebf1^{hcd2+/} Flt3Ltg as specified) were injected intravenously. After 3 weeks, spleen and thymus of recipient mice were separately analysed by flow cytometry.
5.5 Bulk cultures with cytokines

A number of $5 \times 10^4$ sorted haematopoietic progenitors from $Flt3Ltg$ mice were cultured with 50ng/ml Flt3L prepared in-house and 10% IL-7 conditioned medium in a 24-well flat-bottom plate. Cells were maintained as previously indicated and from day 2 to day 6, one well containing cells from each population was analysed by flow cytometry for Ly6D and CD19 expression.

5.6 Statistical analysis

Statistical analysis was performed with GraphPad Prism v6.0f software. Two-tailed unpaired Student $t$ tests were used for statistical comparisons. Data are presented as mean values ± SEM (n.s. not significant or $P > 0.05$, *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$).

5.7 Bulk RNA sequencing

5.7.1 Sample acquisition

Ly6D$^+$ and TN EPLM subpopulations as well as CD115$^-$ TN and Pro-B cells were sorted from femurs of 2-pooled male $Flt3Ltg$ mice (6 to 8 weeks of age). After each sort, cells were centrifuged, resuspended in 0.5ml of TRIzol reagent and stored at -80°C for later total RNA extraction. From the Ly6D$^+$ and TN samples, 100μl containing $\sim3 \times 10^4$ cells were previously used for the capture of single cells.

5.7.2 RNA extraction

Total RNA was extracted from ex-vivo sorted samples using TRIzol-based method [115, 116]. Briefly, $1 \times 10^5$ to $3 \times 10^5$ cells were lysed in 0.5ml of TRIzol reagent and 0.1ml of chloroform was added per 0.5ml TRI reagent. After incubation and centrifugation for phase separation, the aqueous phase containing the RNA was recovered and mixed with isopropanol in a 1:1 ratio for RNA precipitation. Following 15min incubation and centrifugation, the supernatant was discarded while the RNA pellet was first washed with 75% ethanol and subsequently resuspended with 20μl
of DEPC treated water. Concentration and 260/280 purity ratio was initially determined using NanoDrop 1000 Spectrophotometer (Witec AG). Selected RNA samples were stored at -80°C for later usage.

5.7.3 RNA quality
Either 500ng or 1μg of total RNA per sample was sent to the Genomics Facility at the D-BSSE (Basel) for quality control, library preparation and sequencing. Quality and level of degradation of the extracted RNA was assessed with RNA integrity number (RIN) assigned by the Agilent 2100 Bioanalyzer instrument using either the Nano or the Pico Agilent RNA 6000 kit (Agilent Technologies). Samples with a RIN value over 8 and presenting clean peaks were considered for further analysis. The RNA quantity was measured by the Infinite M1000 PRO - Tecan instrument using the Quant-iT RiboGreen RNA Assay Kit.

Figure M1. Quantification of raw sequenced data for the Ly6D⁺, TN and Pro-B samples. (A) Number of sequenced reads per sample. (B) Percentage of reads that mapped to the mouse genome (mm9). (C) Number of counts per sample (library size) considering reads mapped to genes (exons only). (D) The total number of detected genes per sample (with at least 1 count).
5.7.4 Library preparation

For the preparation of sequencing libraries, the TruSeq Stranded mRNA LT Sample Preparation kit was used following the manufacturer’s guide [117]. The polyA containing mRNA molecules were purified using poly-T oligo attached magnetic beads and subsequently fragmented using divalent cations under elevated temperatures. Afterwards, the RNA fragments were copied into first strand cDNA using reverse transcriptase and random primers. Strand-specificity information was achieved by replacing dTTP with dUTP during the second strand cDNA synthesis. To prevent self-ligation of the double-stranded cDNA, the 3’ ends of the blunt fragments were adenylated followed by ligation of barcoded adapters suitable for Illumina-based sequencing. The product was subjected to 15 cycles of PCR amplification.

Size and purity of the library fragments was assessed by the Fragment Analyzer using the NGS Fragment 1-6000bp method (average fragment size 321bp, sd 20.36), while quantification was done with Quant-iT PicoGreen® dsDNA Assay Kit; TEcan instrument.

5.7.5 Sequencing

Indexed DNA libraries were pooled in equal volumes and loaded on one NextSeq 500 High Output flow cell (Illumina). Single-end sequencing was performed on the Illumina NextSeq™ 500 Sequencing System (D-BSSE, Basel) for 81 cycles yielding in 21 to 35 millions of reads, 81-mers, per sample (Fig. M1A).

The Genomics Facility with the Illumina pipeline performed de-multiplexing and reads were transferred in the FastQ format via openBIS. A quality control of the sequenced data was performed using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, Version 0.11.3). All samples comprised high number of reads (> 21 millions) with median Quality Score (QS) of 35, a GC content distribution equivalent to the expected theoretical distribution (~52%), a sequence duplication level typical for RNA-seq samples, and no adapter content present (no need for trimming of reads). Fig. M2 shows a representative example.
Figure M2. Example of quality control of raw sequence data (FASTQC). (A) Quality scores for individual positions within read sequence (over all reads). (B) Quality score distribution over all sequences. (C) GC content distribution over all sequences. Red: theoretical GC content (%); blue: observed GC content (%). (D) Distribution of sequence length over all sequences. (E) Relative number of sequences with different degrees of duplication. (F) Frequency of contamination by sequencing adapters. Replicate 2 of Ly6D group is taken as a representative example.

5.7.6 Pre-processing of sequencing data

All downstream analysis was performed using the open source software R accessed via RStudio server (R version 3.2.0). Sequencing reads were aligned to the mouse genome assembly, version mm9 (downloaded from UCSC http://genome.ucsc.edu), with SpliceMap [118, 119], included in the R/Bioconductor package QuasR, version 1.10.1. Splice-map was also capable of mapping reads that cover exon junctions. More than 80% of total reads were successfully mapped for each sample (Fig. M1 B). Subsequently, a count table with gene expression levels was generated using the qCount function from QuasR package and coordinates of RefSeq mRNA genes (http://genome.ucsc.edu, downloaded in December 2013). The expression level was defined as a number of reads that started within any annotated exon of a gene (exon-union model). Total counts per sample ranged from 13 to 22 millions (Fig. M1 C), the so-called library size. Genes with no counts across all samples were filtered out from the analysis. For 17,290 genes at least 1 read was detected across all samples, corresponding to ~14,800 genes per sample (Fig. M1 D).

Raw counts were normalized between samples with the TMM method (weighted trimmed mean of M-values [120]), expressed as counts per million mapped reads (CPM), and transformed to the log2-scale (log_{2}CPM).
5.7.7 Data analysis

Differential expression analysis was performed using edgeR v3.12.1 [121]. A prior count of 8 was used in order to minimize the large log-fold changes for genes with small number of counts. Genes with a false discovery rate (FDR) <0.05 and \( \text{abs} \mid \log_2(\text{FoldChange}) \mid >1 \) were considered differentially expressed genes (DEG). For principal component analysis (PCA) average gene expression was centred to zero and only the top 50% of genes with highest variance across analysed dataset (calculated as inter-quartile range) were used. PCA plots were generated with the ggplot2 v2.1.0 R package. Heatmap with sample pair-wise Pearson’s correlation coefficients were also generated with the same subset of genes and visualized with the NMF v0.20.6 R package. Annotated heatmap of gene expression variation of the indicated genes across the samples was also generated with the NMF v0.20.6 R package. Average gene expression was centred to 0 and distances were calculated from Pearson’s correlation values of samples using the Euclidean method. MA and Volcano plots were produced using custom R scripts. Gene ontology enrichment analysis was performed with the DAVID v6.7 bioinformatics database, based on Fisher’s Exact method [122, 123]. Gene ontology terms of DEG were determined to be significantly enriched when \( P_{\text{val}} <0.05 \).

5.8 Single-cell RNA sequencing

5.8.1 Capture of single cells

Single cells were captured from ex-vivo sorted haematopoietic progenitors on a small-sized (5-10\( \mu \)m cell diameter) C1 Single-Cell Auto Prep IFC for mRNA sequencing (Fluidigm) using the Fluidigm C1 system. Cell diameter of the Ly6D\(^+\) and TN cells, imaged on Leica DMI 4000 microscope and measured with ImageJ software, was similar and homogeneous, 8.54 and 8.77\( \mu \)m respectively (Fig. M3 A). Therefore, no bias due to cell size or morphology was expected during the capturing. Cells were loaded onto the chip at a concentration of ~300 cells/\( \mu l \) as recommended by the manufacturers and the 96 chambers were inspected by phase-contrast microscopy to determine the number of captured single cells. A total of 3 chips per population were used yielding a good overall capture efficiency, 178 Ly6D\(^+\) and 232 TN single
cells (Fig. M3 B,C). Subsequently, cells were lysed, the polyA containing mRNA molecules were hybridized to oligo-dT and whole-transcriptome full-length amplified cDNA was prepared by template switching on the C1 fluidigm chip according to the manufacturer’s protocol, using the SMARTer Ultra Low RNA kit for the Fluidigm C1 System (Clontech). Quantification of cDNA was done with Quant-iT PicoGreen® dsDNA Assay Kit; TEcan instrument.

5.8.2 Library preparation

Illumina single-cell libraries were constructed in 96-well plates using the Nextera XT DNA Library Preparation Kit (Illumina) following the protocol supplied by Fluidigm (“Using C1 to Generate Single-Cell cDNA Libraries for mRNA Sequencing”). Briefly, 0.1-0.3ng of harvested cDNA was subjected to tagmentation, a process in which the DNA fragmentation and sequencing adapter ligation occurs in a single step performed by the Nextera XT transposome, followed by purification with AMPure XP beads.
5.8.3 **Sequencing**

Indexed DNA libraries originated from single cells captured in 3 different chips (288 libraries) were pooled in equal volumes and loaded on one NextSeq 500 High Output flow cell (Illumina). Single-end sequencing was performed on the Illumina NextSeq™ 500 Sequencing System (D-BSSE, Basel) for 76 cycles.

As for the bulk RNA sequencing, reads (76-mers) were received after de-multiplexing in the FastQ format and checked for quality using the FastQC tool ([http://www.bioinformatics.babraham.ac.uk/projects/fastqc/](http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), Version 0.11.3). Only the FastQ files corresponding to C1 chambers with a single cell were selected for downstream analysis, thus excluding the ones derived from doublets, debris or empty chambers (previously determined by phase-contrast microscopy). We obtained a total of 360 million reads for the Ly6D^+ and 371 million reads for the TN. The average number of reads per cell was 2x10^6 for the Ly6D^+ and 1.6x10^6 for the TN (Fig. M4 A).

![Figure M4. Quantification and quality control of raw sequenced data for Ly6D^+ (left panels) and TN (right panels) single cells. Per cell distribution of (A) number of sequenced reads, (B) percentage of reads mapped to the mouse genome (mm9) out of the total number of reads, (C) counts (library size) considering reads mapped to genes (exons only) and, (D) total number of genes detected (with at least 1 count). Ly6D^+ n=178; TN n=213. Blue line: mean; dotted red line: thresholds applied to the data (any cells not meeting these thresholds failed the quality control and were excluded from the analysis).](image)
5.8.4 Pre-processing of sequencing data

All downstream analysis was performed using the open source R software accessed via RStudio server (R version 3.2.0). Sequencing reads were aligned to the mouse genome assembly, version mm9 (downloaded from UCSC http://genome.ucsc.edu), with SpliceMap [118, 119], included in the R/Bioconductor package QuasR, version 1.10.1. Approximately 80% of total reads were successfully mapped for each sample (Fig. M4 B). Subsequently, a count table with gene expression levels was generated using the qCount function from QuasR package and coordinates of RefSeq mRNA genes (http://genome.ucsc.edu, downloaded in December 2013). The expression level was defined as the number of reads that started within any annotated exon of a gene (exon-union model). Total counts per cell were approximately 810,000 for the Ly6D+ and 720,000 for the TN (Fig. M4 C), the so-called library size. Genes with no counts across all samples were filtered out from the analysis. At least one read per gene was detected for a total of 14,814 genes across all cells, corresponding to approximately 3,500 expressed genes per cell in both Ly6D+ and TN (Fig. M4 D).

During the quality control, cells having less than 60% of mapped reads, less than 200,000 counts, or less than 800 detected genes were filtered out from further analysis (dotted red lines in Fig. M4). In total, 365 (152 Ly6D+ and 213 TN) cells out of 410 or 89% passed these criteria.

Raw counts were normalized between cells and genes, expressed as fragments per kilobase of transcript per million mapped reads (FPKM). For visualization purposes, 1 was added to FPKM values and transformed to the log2-scale (log2(FPKM)).

5.8.5 Data analysis

If not otherwise specified, the downstream analysis was performed using the 1008 differentially expressed genes (DEG, FDR <0.05 and abs|log2(FoldChange)| >1) from the bulk RNA-seq experiment when comparing Ly6D+ with TN populations.

Dimensionality reduction was performed with principal component analysis. Average gene expression was centred to zero and PCA plots were generated with the ggplot2 v2.1.0 R package. To visualize the degree of cell-to-cell heterogeneity,
annotated heatmap of sample pair-wise Pearson’s correlation coefficients was produced using the NMF v0.20.6 R package. Eight Ly6D+ cells were not considered for subsequent clustering because of their very low transcriptome correlation to any other cell, on average less than 0.3 (Fig. 11C left, in results section). Cell clustering was performed using the Partitioning Around Medoids (PAM) method implemented in the cluster v2.0.4 R package [124]. Gene expression was first centered (mean=0) and distances were calculated from cell-to-cell Pearson’s correlation values using the Euclidean method. The optimal number of clusters was selected based on silhouette plot, which for Ly6D+ corresponded to K=2 (with average silhouette width of 0.10) and K=3 for the TN (with average silhouette width of 0.13). Cells with negative silhouette width values were excluded while the other 331 cells were assigned to one of the 5 groups. Heatmap with Pearson’s correlation coefficients among the clustered groups of cells was generated with the top 50% of genes with highest variance across analysed dataset (calculated as inter-quartile range) and visualized with the NMF v0.20.6 R package.

Differential gene expression analysis to compare the clustered groups of cells was performed using the 14,528 detected expressed genes across the 331 single cells with edgeR v3.12.1 [121]. A prior count of 0.5 was added to all gene counts in order to minimize the large log-fold changes for genes with small number of counts. Genes with FDR <0.05 and abs|log2(FoldChange)| >1 were considered as differentially expressed. MA, Volcano, Violin and Scatter plots were produced using custom R scripts.

Gene ontology enrichment analysis was performed with the DAVID v6.7 bioinformatics database, based on Fisher’s Exact method [122, 123]. Gene ontology terms of DEGs were determined to be significantly enriched when FDR <0.05.
6. Results

6.1 EPLM progenitor population can be divided into at least 4 subpopulations

We have previously characterized an uncommitted and multipotent haematopoietic B220⁺ CD117_{int} CD19⁻ NK1.1⁻ progenitor with combined lymphoid and myeloid potential that we have called “Early Progenitor with Lymphoid and Myeloid potential” (EPLM) [33]. When describing progenitor cells that apparently have multiple lineage potentials, there is a continuing debate concerning whether individual cells within the seemingly phenotypically homogeneous population are truly multipotent or whether the population contains a mixture of cells each with different lineage potentials. With the aim of addressing this question, we decided to examine the expression of additional markers known to be associated with different haematopoietic lineages. These markers were Ly6D, SiglecH and CD11c. Ly6D (Lymphocyte Antigen 6 Complex, Locus D) is a receptor with unknown function used as a specification marker for early B-cell stages. Indeed, a Ly6D⁺ subpopulation of CLP has already been shown to have a B-cell bias [125, 126]. SiglecH is an immunoglobulin-like lectin receptor expressed on mouse plasmacytoid dendritic cells (pDC) that upon binding to sialic acid carbohydrates modulates the secretion of type I interferons and is commonly used as a pDC marker [127, 128]. CD11c (integrin alpha-X) associates with CD18 (integrin beta-2 chain) forming the complex CD11c/CD18 that is the receptor for the complement component iC3b as well as for fibrinogen and is also involved in cell-cell interactions during inflammatory responses. It is important for monocyte adhesion and is used as a murine marker for conventional dendritic cells (cDC) [129].

To test the expression of the selected markers, we combined the reported EPLM gating strategy to define B220⁺ CD117_{int} CD19⁻ NK1.1⁻ EPLM and included Ly6D and SiglecH. This staining combination resulted in the identification of three EPLM fractions (F) both in WT and Flt3Ltg mice. Thus, F1 was Ly6D⁺ SiglecH⁻, F2 Ly6D⁺ SiglecH⁺, and F3 Ly6D⁻ SiglecH⁻ (Fig. 4A,B upper panels).
Figure 4. Heterogeneous expression of Ly6D, SiglecH and CD11c on EPLM progenitor population (B220^+ CD117^int CD19^- NK1.1^-). (A,B) Representative FACS plots of EPLM from the BM (2 femurs) of WT (A) and Flt3Ltg mice (B) with the addition of Ly6D and SiglecH identifying three fractions (illustrated as F1, F2 and F3). CD11c expression in each of the fractions is shown in the second row. (C) Percentage of CD11c^+ cells in each EPLM fraction from WT (n=5) and Flt3Ltg (n=5) mice. Bars show mean±SEM. (D) Representative FACS plot showing the gating strategy for the four EPLM subpopulations from WT (upper panel) and Flt3Ltg (lower panel) mice. (E) Summary of EPLM subpopulations as percentages (upper panel) or numbers (lower panel) from WT (n=5) and Flt3Ltg (n=5) mice. Shown as mean±SEM. Two-tailed unpaired Student t tests, *P ≤ 0.05, **P ≤ 0.01.
Subsequently, we explored CD11c expression by these three fractions in both mouse strains. Results indicated that F1 contained only 5.1±0.5% CD11c+ cells whereas F2 were mostly CD11c+ (85.2±2.7%). Interestingly, F3 were heterogeneous for CD11c expression, with about one third (28.1±2.7%) being CD11c+ in WT but considerably more (43.1±9.1%) in Flt3Ltg mice (Fig. 4A, B lower panels and Fig. 4C). This result indicates that F3 can be further subdivided into two CD11c+ and CD11c− fractions resulting in four major EPLM subpopulations. In Fig. 4D and in subsequent experiments, we represent the 4 EPLM subpopulations in a simplified manner by staining for Ly6D in one colour and both SiglecH and CD11c using antibodies conjugated with the same fluorochromes in the other. We can thus distinguish the four EPLM subpopulations as: the Ly6D+ SiglecH− CD11c− (Ly6D+) single positive subpopulation in red, the Ly6D− SiglecH− CD11c− triple negative (TN) in green, the Ly6D+ SiglecH+ CD11c+ triple positive in blue. Since this is the only subpopulation that contains significant numbers of SiglecH+ cells (Fig. 4A, B upper right cytogram), they are referred to as SiglecH+. Finally, the Ly6D− SiglecH− CD11c+ (CD11c+) subpopulation is shown in orange (Fig. 4D, E). In the bone marrow of WT mice, EPLM subpopulations are present in limited numbers (2.7±0.8x10^3 Ly6D+ cells, 2.3±0.7x10^3 TN cells, 3±0.3x10^3 SiglecH+ cells, and only 1.1±0.3x10^3 CD11c+ cells). In marked contrast, Flt3Ltg mice contain two orders of magnitude more of each EPLM subpopulation (2.4±0.9x10^5 Ly6D+, 0.7±0.2x10^5 TN, 1.8±1x10^5 SiglecH+, and only 0.5±0.2x10^5 CD11c−) (Fig. 4E lower panel) with minor modifications in their relative frequencies (Fig. 4E upper panel). We have previously shown [109] that Flt3Ltg mice had dramatically increased numbers of EPLM and this result shows that all four subpopulations are affected almost equally.

In conclusion, based on the expression of Ly6D, SiglecH and CD11c, the EPLM progenitor population can be further subdivided into at least four subpopulations whose numbers all increase significantly in Flt3Ltg mice.
6.2 The heterogeneity of EPLM is reflected in different sets of potentials for each subpopulation

We next wanted to assess if the heterogeneous expression of Ly6D, SiglecH and CD11c cell surface markers by EPLM reflects distinct developmental potentials. Several clonal in vitro culture systems exist to quantify the capacity of progenitor cells to differentiate into distinct lineages. These systems use stromal cells and addition of exogenous factors known to support differentiation towards a particular lineage. Stromal cell clones supporting B lymphopoiesis and/or myelopoiesis are well established and extensively used. ST2 stromal cells producing macrophage colony-stimulating factor (M-CSF) support both B lymphopoiesis and/or myelopoiesis although the latter one is favoured [112, 130]. B-cell development from haematopoietic progenitors is much more efficient on OP9 stromal cells [113, 131] because they are derived from the bone marrow of M-CSF deficient op/op mice.

Figure 5. Distinct in vitro developmental potentials for the EPLM subpopulations. Limiting dilution analysis of Ly6D⁺, TN, SiglecH⁺ and CD11c⁺ for B-cell (A), T-cell (B) or myeloid (C) potentials. EPLM subpopulations were sorted from WT (10 pooled mice, upper panels) and Flt3Ltg (2 pooled mice, lower panels) and plated at the indicated concentrations on either OP9 stromal cells together with IL-7 (A), OP9-DL1 stromal cells in the presence of IL-7 (B), or ST2 stromal cells (C). After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1 and ST2 cell cultures), B-cell, T-cell or myeloid clones were scored using an inverted microscope and, when unclear, clone type was confirmed by flow cytometry. TN and CD11c⁺ subsets from WT were plated at only 3 cell concentrations due to the limited number of cells. The lowest concentration of the two previous subsets was plated in 96 replicates while all the rest in 48. One representative experiment is shown and Table 2 summarizes the independent experiments performed and indicates the n repetitions for each subpopulation, mice and condition.
In the absence of M-CSF, macrophage growth is not supported thereby revealing lymphocyte potential [130]. This culture system is normally supplemented with IL-7 and/or other cytokines promoting cell growth such as Flt3L or SCF. Even though B-cell and myeloid propagation was for a long time successfully achieved in monolayer stromal cell cultures, a T-cell system had not been reported. Then, in 2002 Zuniga-Pflücker and colleagues showed for the first time the efficient and long-term commitment and propagation of T cells in a simple stromal cell monolayer system [114, 133]. Use was made of the well-established OP9 stromal cell line, but in addition, since the Delta-Notch system had been identified as instructing T-cell lineage choice [134], OP9 cells were transduced with the Notch-ligand Delta-like 1 (DL1). This generated the OP9-DL1 stromal cell line, which efficiently promoted early stages of T-cell development.

We made use of the above described culture systems and performed a limiting dilution (LD) assay in order to examine the B-cell, T-cell and myeloid precursor frequencies of the four EPLM subpopulations. For that, we sorted them from WT or Flt3Ltg mice and plated graded number of cells with the appropriate cocktail of stromal cells and cytokines to support B-cell (OP9 + IL7), T-cell (OP9-DL1 + IL7) or myeloid (ST2) differentiation. After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1 and ST2 cell cultures), B-cell, T-cell or myeloid clones were scored using an inverted microscope and, when unclear, clone type was confirmed phenotypically by flow cytometry. The result of a representative experiment is shown in Fig. 5 and a summary of all experiments in Table 2. Under B cell conditions, Ly6D+ and TN

<table>
<thead>
<tr>
<th></th>
<th>WT</th>
<th>Flt3Ltg</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B-cell potential</td>
<td>T-cell potential</td>
<td>Myeloid potential</td>
</tr>
<tr>
<td>Ly6D+ n</td>
<td>11</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>mean freq</td>
<td>1 in 6.4</td>
<td>1 in 38 &lt;1 in 99</td>
</tr>
<tr>
<td>SiglecH+ n</td>
<td>4</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>mean freq</td>
<td>&lt;1 in 500</td>
<td>&lt;1 in 500</td>
</tr>
<tr>
<td>TN n</td>
<td>7</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>mean freq</td>
<td>1 in 34</td>
<td>1 in 55</td>
</tr>
<tr>
<td>CD11c+ n</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>mean freq</td>
<td>&lt;1 in 500</td>
<td>&lt;1 in 500</td>
</tr>
</tbody>
</table>

Table 2. Summary table containing mean frequencies of progenitors with B-cell, T-cell and myeloid potential. Number of independent experiments is shown (n) for the different populations.
subpopulations generated colonies, thus revealing B-cell potential. The B-cell precursor frequency was higher in the Ly6D\(^+\) cells (about 1 in 5) than in the TN (about 1 in 20) and overall greater for both subpopulations in WT mice compared with their respective counterparts in Flt3Ltg mice (1 in 12 for Ly6D\(^+\) and about 1 in 90 for TN) (Fig. 5A). CD11c and SiglecH\(^+\) subpopulations did not generate colonies either under B-cell or T-cell conditions (Fig. 5A,B). Ly6D\(^+\) and TN cells from WT mice generated T cells at low frequencies whereas in Flt3Ltg mice, frequencies were notably increased (Fig. 5B). When EPLM subpopulations were plated on ST2 stromal cells, the Ly6D\(^+\) subpopulation did not generate myeloid clones, whereas all other EPLM subpopulations, from either WT or Flt3Ltg mice, possessed myeloid potentials although at different frequencies (Fig. 5C). Therefore, each EPLM subpopulation has different sets of potentials. Whereas Ly6D\(^+\) cells are lymphoid restricted, TN show trilineage developmental potential. The SiglecH\(^+\) and CD11c\(^+\) subpopulations do not show lymphoid potential and as their cell-surface marker profile suggests, they could be (at least a fraction of them) already committed to the pDC and cDC fates respectively. Since in our laboratory we are mostly interested in studying the lymphoid development of EPLM, for the rest of the project we focused on the Ly6D\(^+\) and TN subpopulations.

We therefore tested the in vivo capacity of Ly6D\(^+\) and TN progenitors to reconstitute the B- and T-cell compartments of lymphopenic mice. The two EPLM subpopulations were sorted from WT or Flt3Ltg mice and equal numbers (4x10\(^3\) from WT and 2x10\(^4\) from Flt3Ltg) transferred into sub-lethally irradiated Rag2-deficient recipient mice [110]. Rag2-deficient mice lack mature B and T lymphocytes due to their inability to initiate V(D)J rearrangement. Therefore, the developmental stages observed after antigen receptor rearrangement are generated from donor cells. Flow cytometry of the spleen at 3 weeks following transfer revealed that B-cell compartments were significantly reconstituted in all mice. Both Ly6D\(^+\) and TN progenitors either from WT or Flt3Ltg generated CD19\(^+\) IgM\(^+\) B cells (Fig. 6A,C). Additional analysis of the spleen CD19\(^+\) cells revealed the presence of both CD21\(^\text{high}\) CD23\(^-\) marginal zone B cells (MZB) and CD21\(^\text{int}\) CD23\(^+\) follicular B cells (FB) (Fig. 6A lower panels). Therefore, although TN cells present less B-cell precursor frequency in vitro, both populations have in vivo B-cell developmental potential. Analysis of the
Figure 6. Reconstitution of B- and T-cell compartments in sub-lethally irradiated B6 Rag2-deficient mice with Ly6D⁺ or TN cells from WT or Flt3Ltg mice. (A) CD19 versus IgM expression on spleen cells 3 weeks after transfer of $4 \times 10^3$ WT derived (upper panels) or $2 \times 10^4$ Flt3Ltg derived (lower panels) Ly6D⁺ and TN subpopulations. Also CD21 and CD23 expression on gated CD19⁺ spleen cells are shown. (B) CD4 and CD8 expression on thymocytes 3 weeks after transfer of $4 \times 10^3$ WT derived (upper panels) or $2 \times 10^4$ Flt3Ltg derived (lower panels) Ly6D⁺ and TN subpopulations. Also TCRβ expression on DN, DP, CD8⁺ and CD4⁺ gated thymocytes is shown. (C,D) Quantification of CD19⁺ IgM⁺ (C), CD4⁺ and CD8⁺ (D) populations presented as frequency of live cells. Ly6D⁺ WT (n=5), TN WT (n=4), Ly6D⁺ Flt3Ltg (n=5), TN Flt3Ltg (n=5). Shown as mean±SEM.
thymus showed that, in line with the in vitro observations, cells derived from Flt3Ltg were much more efficient at reconstituting the thymus than those from WT mice (Fig. 6B,D). Thus, WT derived progenitors could only reconstitute the thymus of one out of nine mice whereas progenitors from Flt3Ltg mice reconstituted seven out of ten animals (Fig. 6D). Comparing the two EPLM subpopulations, in WT mice only the Ly6D⁺ subpopulation had any T cell reconstitution potential (1/5 vs 0/5) whereas in Flt3Ltg mice, both subpopulations had quite robust reconstitution potential (4/5 vs 3/5) (Fig. 6D). In the T-cell reconstituted thymus, we observed normal levels of CD4⁺ CD8⁺ double positive (DP), and CD4⁺ or CD8⁺ single positive T cells. Simultaneous staining for TCRβ indicated that, like in the normal thymus, TCRβ expression on DP was lower compared to that on CD4⁺ cells, and that the CD8⁺ population contained a mixture of TCRβ⁺ and TCRβlow “immature” CD8⁺ cells (Fig. 6B). Moreover, CD4⁻ CD8⁻ double negative (DN) thymocytes were largely TCRβ⁻ and only a small fraction was TCRβ⁺. Interestingly, although Ly6D⁺ Flt3Ltg cells showed higher in vitro T cell potential (Fig. 6A,B lower panels), they reconstituted better the B-cell than the T-cell compartment.

6.3 Ly6D⁺ EPLM has a lymphoid whereas TN a myeloid genetic signature

In order to characterize EPLM subpopulations at the molecular level, we performed gene expression profiling (RNA-seq) of bulk populations of Ly6D⁺ and TN EPLM subpopulations as well as CD117⁺ CD19⁺ Pro-B cells from Flt3Ltg mice. We included the latter population as an already B-cell lineage committed bone marrow population and thus downstream of EPLMs. Ly6D⁺ and TN were sorted as indicated in Fig. 4D. All populations were sorted four separate times from the pooled bone marrow of two 6-8 week-old male Flt3Ltg mice, in which EPLM numbers are significantly increased. All samples were processed as explained in Materials and Methods and all passed the quality control (Fig. M1 and Fig. M2). Moreover, all biological replicates presented high transcriptome correlation (> r=0.990) (Fig. 7). Therefore, we proceeded to apply Principal Component Analysis (PCA) to the data
and calculated sample pair-wise Pearson’s correlations. The highest variation among samples (71.13% reflected in PC1 axis) accounted for their developmental stage; with the uncommitted EPLM subpopulation on the left and the committed Pro-B cells on the right (Fig. 8A). Pair-wise transcriptome correlation revealed that, in line with PCA, Ly6D⁺ and TN were the two populations with the highest transcriptome association ($r=0.973$) whereas from the two EPLM subpopulations, the Ly6D⁺ subset was the closest to Pro-B cells ($r=0.918$ Ly6D⁺/Pro-B versus $r=0.886$ TN/Pro-B) and therefore to the B-cell lineage (Fig. 8B).

We further explored transcriptome differences by performing differential expression analysis and this is summarised in the table in Fig. 8C showing the total number of Differentially Expressed Genes (DEG) for each pair-wise comparison as well as the number and the fraction corresponding to up-regulated and down-regulated genes. To be considered as DEG, the gene expression had to be at least two times higher or lower in one population compared with the other and this change in expression had to be significant (corrected p-value <0.05). A considerable
fraction (21% and 17%) of genes was differentially expressed when comparing Pro-B with either Ly6D⁺ or TN subpopulations respectively whereas only about 6% (1008 genes) presented a significant change in expression between the two EPLM subpopulations (Fig. 8C), thus suggesting again that Ly6D⁺ and TN cells are more related to each other than to Pro-B cells. In addition, the graphical representation of the differential expression analysis (volcano plots, Fig. 8D) also revealed that the fold changes and significances of DEGs in the Ly6D⁺ vs TN was less (Fig. 8D right panel).
than when they were individually compared with Pro-B (Fig. 8D left and middle panels). Interestingly, there was a similar fraction of up-regulated and down-regulated genes in each comparison, indicating no predominant activation or repression of genetic programmes from one haematopoeietic stage to the other (Fig. 8C).

We next investigated in detail the genetic differences between the apparently transcriptomically related Ly6D+ and TN EPLM subpopulations. For this, we studied the nature of the DEG (Ly6D+ vs TN) by gene ontology (GO) enrichment analysis. This analysis revealed that the 500 up-regulated genes in Ly6D+ cells were enriched for lymphoid biological processes such as activation, proliferation and differentiation of B and T cells, VDJ recombination and immunoglobulin production...

Table 3. Selection of enriched Biological Processes (eBP) in up-regulated and down-regulated genes of the Ly6D+ versus TN comparison. Complete list in excel file 2 (2_eBP_Ly6DvsTN).

We next investigated in detail the genetic differences between the apparently transcriptomically related Ly6D+ and TN EPLM subpopulations. For this, we studied the nature of the DEG (Ly6D+ vs TN) by gene ontology (GO) enrichment analysis. This analysis revealed that the 500 up-regulated genes in Ly6D+ cells were enriched for lymphoid biological processes such as activation, proliferation and differentiation of B and T cells, VDJ recombination and immunoglobulin production... (Table 3, up). Moreover, the genes that accounted for B-cell related biological
processes were highly expressed and most of them within the top 50 of up-regulated genes (Fig. 9A, upper panel). Among these, were genes encoding the B-cell related transcription factors Pax5, Ebf1 and Pou2af1 (Obf1), the recombinase machinery.
Rag1 and Rag2, the surrogate light chains VpreB1, VpreB2, VpreB3 and Igl1 (lambda 5) of the pre B-cell receptor (pre-BCR), the signalling immunoglobulin α (Cd79a) and β (Cd79b) chains of the pre-BCR complex, the non-receptor tyrosine kinase Blk involved in BCR signaling, the receptor for interleukin-7 Il7r and other lymphoid related genes (Dntt, Lax and, as expected, Ly6d itself). Interestingly, although Ly6D+ cells were sorted as CD19- cells, mRNA expression of the B cell co-receptor CD19 was already detected. Taken together, these results suggest that qualitatively, Ly6D+ cells express a B-cell genetic signature characteristic of CD19+ Pro-B cell stage. However, quantitatively, the overall expression of these genes is markedly lower than Pro-B cells (Fig. 9B). In contrast to B-cell related genes, the genes accounting for T-cell related biological processes in the 500 up-regulated genes in Ly6D+ cells presented lower expression ratios and variable expression intensities overall (Fig. 9A, middle panel). Among these genes were the T-cell transcription factor Bcl11b, the Notch1 receptor, a master regulator of T-cell commitment whose signaling represses the expression of genes related with other lineages, the signalling CD3 zeta chain (Cd247) of the T-cell receptor (TCR) complex, genes involved in pre-TCR signalling such as Lck (non-receptor tyrosine kinase) Rhoh (related GTP-binding protein) Zap70 (tyrosine kinase) and Sla2 (Src-like-adapter protein), Trat1 (an adaptor protein that stabilizes the TCR/CD3 complex at the surface of T-cells), the tnfsf11 cytokine and Nlrc3 (positive regulators of T cell activation), and the inhibitory T-cell related receptors Ctica4 and Ctica2b. This T-cell genetic signature is exclusive to the Ly6D+ subpopulation (Fig. 9B) and is the “feature” that separates them from both the Pro-B and TN cells along the PC2 of the principal component analysis (Fig. 8A).

Analysis of the 500 down-regulated genes revealed that they were largely related with myeloid and innate biological processes such as inflammation, phagocytosis, responses to bacteria, yeast and fungi, and macrophage activation (Table 3, down). Some key genes accounting for these processes were Mpo, Elane and Ctsg (enzymes with microbicidal activity), Prtn3 (a serine protease that degrades elastin, fibronectin, laminin, vitronectin, and collagen), the phagocytic Fc receptors Fcgr2b, Fcgrt, Fcerg1g and Fcgr1, the pathogen recognition receptor Tlr1 (Toll-like receptor 1), Gata2 (transcriptional regulator of phagocytosis), Clec7a (involved in TLR2-mediated inflammatory responses), the polysaccharide binding protein Lbp and
the chemokine receptor Cx3cr1 involved in myeloid leukocyte activation (Fig. 9A, lower panel, and B). Therefore, TN cells would appear to present a myeloid genetic signature.

From this transcriptomic analysis, we conclude that i) EPLM subpopulations are distinct from one another and are both distinct from Pro-B cells ii) of the two EPLM subsets, Ly6D+ cells are closer to Pro-B cells iii) whereas the Ly6D+ subset has a largely lymphoid genetic signature, that of the TN subset is more myeloid.

6.4 EPLM subpopulations are developmentally related

We have seen that Ly6D+ and TN cells have distinct genetic signatures and different in vitro developmental potentials. However, they are fractions of the original EPLM population. The question arose whether as a population EPLM are composed of developmentally unrelated fractions or whether there is a precursor-product relationship between them. To address this question, we sorted the two EPLM subpopulations from Flt3Ltg mice and cultured them in the presence of Flt3L and IL-7. An initial number of 5x10^4 cells were plated and from day 2 to day 6 we monitored the emergence of Ly6D+ and CD19+ cells by flow cytometry. We observed that, already after 48 hours, almost half of the cells sorted as Ly6D negative (TN), up-regulated Ly6D (Fig. 10A lower left panel). This indicates that the TN EPLM...
subpopulation can give rise to the Ly6D⁺ subset. B cells expressing CD19 were detected around day 4 reaching 24% of cells at day 6, whereas by this time, Ly6D expression had decreased (Fig. 10A lower panels). In contrast, Ly6D⁺ cells had differentiated into B cells already by day 2 and 75% were CD19⁺ at day 6 (Fig. 10A upper panels). Moreover, the total number of B cells generated from Ly6D⁺ was dramatically higher than those generated from TN cells (Fig. 10B). These results indicated that, as observed in the limiting dilution assays, Ly6D⁺ cells have higher B-cell precursor frequency and differentiate into B cells with faster kinetics compared with TN. The slower kinetics observed in the TN cells indicate that the TN are the direct precursors of the Ly6D⁺ and, therefore, they might first move to the Ly6D⁺ stage before differentiating into B cells.

6.5 Single-cell RNA sequencing reveals distinct degree of molecular heterogeneity of Ly6D⁺ and TN EPLM subpopulations

As a population, the TN subset of EPLM would appear to have multilineage developmental potential (Fig. 5) prompting the question whether they are composed of a mixture of cells. If this were the case, which fraction of TN EPLM constitutes the precursor of the Ly6D⁺ fraction? Therefore, we further explored the heterogeneity of EPLM subpopulations by performing single-cell RNA sequencing. Ly6D⁺ (B220⁺ CD117int CD19⁻ NK1.1⁻ Ly6D⁺ SiglecH⁻ CD11c⁻) and TN (B220⁺ CD117int CD19⁻ NK1.1⁻ Ly6D⁻ SiglecH⁻ CD11c⁻) were sorted from the same Flt3Ltg mice used for the bulk RNA sequencing and utilized for the capture of single cells with the C1 Fluidigm system. A total of 178 Ly6D⁺ and 232 TN single cells were captured (Fig. M3 C) and their subsequent library preparation and sequencing were performed as explained in Materials and Methods. During the quality control of the sequenced data only cells with more than 60% of mapped reads, at least 2x10⁵ counts and more than 800 detected genes were selected for further analysis, resulting in 152 Ly6D⁺ and 213 TN single cells (Fig. M4). In Principal Component Analysis cells did not cluster according to the chip they were captured, thus revealing that there was no batch effect (Fig. 11A). Instead, in the same PCA with the cells coloured according
to their cell type, we found that the first component (PC1 axis) partially segregated the TN (on the left) from the Ly6D\(^+\) (on the right) with some cells overlapping in the middle (Fig. 11B and attached 3D_PCAplot_2), suggesting that the two EPLM subpopulations, or at least a fraction, are transcriptionally related. PC2 (Fig. 11B) and PC3 (attached 3D_PCAplot_2) showed that Ly6D\(^+\) cells are distributed along a single pathway that is relatively homogeneous whereas the TN cells are more differently distributed and exhibit a branching pattern, thereby indicating that they might be more heterogeneous. In order to better address this, we quantified the degree of cell-to-cell heterogeneity by calculating the cell pair-wise Pearson’s correlation coefficients and generating a correlation heatmap for each cell type (Fig. 11C). Ly6D\(^+\) single cells showed an overall higher transcriptome correlation (predominant yellow/orange colour and a mean correlation value of 0.42, Fig. 11C left) compared with the TN single cells (predominant blue colour and a mean correlation value of 0.32, Fig. 11C right).
0.32, **Fig. 11C** right). Moreover, Ly6D\(^+\) cells showed a seemingly homogeneous correlation whereas the TN cells presented clusters of cells transcriptomically related to each other (yellow) but very different to the rest (blue). These results indicate that the TN EPLM subpopulation has in turn a heterogeneous transcriptome and might be composed by a mixture of cells.

### 6.6 Identification of two Ly6D\(^+\) and three TN subgroups with distinct genetic signatures

The next question after identifying a heterogeneous subset was to unravel how many subgroups did it contain. Therefore, we performed cell clustering using the Partioning Around Medoids (PAM) method as explained in Materials and Methods. The analysis revealed two robust groups (G1 and G2) of Ly6D\(^+\) and three (G1, G2 and G3) of TN cells as illustrated in the PCA plot coloured according to the subgroups (**Fig. 12A** and attached **3D_PCAplot_3**). Thus, the Ly6D\(^+\) population is further subdivided into G1 Ly6D\(^+\), composed of 56 cells (red) and G2 Ly6D\(^+\), composed of 82 cells (orange), whereas the TN population is subdivided into G1 TN with 85 cells (purple), G2 TN composed of 52 cells (blue) and G3 TN with 56 cells (green). In order to explore the degree of similarity or divergence among the clustered groups of cells we calculated their pair-wise transcriptome correlation, which revealed that the two Ly6D\(^+\) subgroups had higher transcriptome association (R=0.696) than those observed between any of the TN subgroups (R=0.582 G1/G2, R=0.662 G1/G3, R=0.666 G2/G3) (**Fig. 12B**). Interestingly, the two groups of cells with the highest transcriptome correlation are part of different EPLM subpopulations, namely the G2 Ly6D\(^+\) and G1 TN with R=0.721 (**Fig. 12B**), thereby suggesting that the G1 TN (purple cells in **Fig. 12A**) group could be the fraction of the TN population that, as we have observed in culture, is the precursor of the Ly6D\(^+\) cells (circles in **Fig. 12A**). Finally, the subgroup having the most distinct transcriptome profile is the G2 TN (**Fig. 12B** left column, and blue cells in **Fig.12A**).

We further studied genetic differences by applying differential expression analysis to the clustered groups of cells. The table in **Fig. 12C** shows all the
comparisons analysed reporting the number and percentage of DEG (FDR <0.05 and abs|log₂(FoldChange)| >1), up-regulated and down-regulated genes. The complete table containing the DEG for each comparison is provided as a supplementary file (3_DEGlists_scRNAseq). Less than 1% of the genes were differentially expressed when comparing the two Ly6D⁺ subgroups (Fig. 12C first row) and with overall low significance level (Fig. 12D, left). In contrast, comparisons among the TN subgroups
yielded more differentially expressed genes (1-6%, **Fig. 12C** second to fourth rows), and with overall higher significance levels and fold ratios (**Fig. 12D**). Of note, when comparing the two subgroups with the highest transcriptome correlation, namely G2 Ly6D⁺ and G1 TN, only 25 genes were differentially expressed with Ly6d, as expected, on the top (complete DEG lists in *3_DEGlists_scRNAseq* excel file), therefore confirming that these two subgroups of cells sorted as two phenotypically distinct EPLM subpopulations (Ly6D⁺ and Ly6D⁻) are related.

We next performed a detailed screening based on the DEG lists in order to unravel the gene expression patterns of each subgroup. **Fig. 13 (A-C)** shows a collection of genes exclusively or more highly expressed in one of the clustered groups of cells. The G1 Ly6D⁺ red subgroup up-regulates genes (**Fig. 13A**) related with B-cell biological processes (**Table 4A**). Although we have previously shown that the entire Ly6D⁺ population is lymphoid restricted (**Fig. 5**) and has a strong B-cell genetic signature (**Fig. 9A** upper panel), single-cell transcriptomic analysis reveals that this signature is mostly contained within the red subgroup, as is exemplified by CD19 expression (**Fig. 13A**), a hallmark of B-lymphopoiesis. The G2 TN blue cluster of cells up-regulates genes characteristic of the cDC lineage such as the genes encoding the α and β chains of the major histocompatibility complex (MHC) class II (H2-Aa, H2-Ab1, H2-Eb1), the MHC class II-associated invariant chain (Ii, Cd74), the transcriptional co-activator of MHC class II genes (Ciita), the E3 ligase (March1) that ubiquitinates the cytoplasmic tail of MHC class II regulating their steady-state level of expression, the transcriptional regulators Id2 and Batf3 (expressed by cDC and necessary for the development of CD8α cDC) and the integrin Itgb7 related with leukocyte migration to mucosal tissues (**Fig. 13B**). Apart from antigen processing and presentation biological processes, the genes up-regulated in the G2 TN cells are also involved in actin cytoskeleton organization, leukocyte adhesion, actin polymerizations and depolymerisation, protein complex assembly, and regulation of cellular component size (**Table 4B** and lists of DEG (G1vsG2)TN and (G2vsG3)TN in excel file *3_DEGlists_scRNAseq*). This suggests that the blue group of clustered cells, which is the most transcriptionally different to the rest (**Fig. 12**), might already be expressing the intracellular machinery necessary to acquire the dendritic cell morphology and the antigen presenting function characteristic of mature cDC. The
G3 TN green subgroup mostly up-regulates myeloid related genes (Fig. 13C) involved in innate biological processes (Table 4C) characteristic of the entire TN population (Fig. 9A lower panel), thus suggesting that this might be the fraction of TN largely containing the observed myeloid potential. However, although the G3 TN subset is myeloid biased, its genetic signature is not resolved to any specific myelomonocytic cell. Some genes such as myeloperoxidase (Mpo) are characteristic of both monocyte/macrophages and granulocytes while others are exclusive to granulocytes (Elane, expressed in neutrophils) or monocyte/macrophages (chemokine Cx3cr1).
Interestingly, our screening did not find genes exclusively expressed in G2 Ly6D<sup>+</sup> orange cells or G1 TN purple cells. Instead, we observed that these two subsets follow the same pattern of expression, meaning that when a gene is expressed in one group it is also expressed in the other one (Fig. 13D) or it is down-

### Table 4. Selection of enriched Biological Processes (eBP) in up-regulated genes of the G1 Ly6D<sup>+</sup> (A), G2 TN (B), G3 TN (C), G1 TN (D) groups of cells compared with the subgroups indicated in brackets. Complete list in excel file (4_eBP_scRNAseq_subgroups).

#### G1 Ly6D<sup>+</sup> (compared with G2 Ly6D<sup>+</sup>)

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Count</th>
<th>PValue</th>
<th>Fold Enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0050853:8-ribosome, RNA syntheses, RNA processing, RNA catabolism</td>
<td>3</td>
<td>2.16E-04</td>
<td>107.84</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0050851:8-antigen receptor-mediated signaling pathway</td>
<td>3</td>
<td>0.0023119</td>
<td>41.94</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0042113:8-cell activation</td>
<td>3</td>
<td>0.00967014</td>
<td>19.36</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:002253:activation of immune response</td>
<td>3</td>
<td>0.0116608</td>
<td>17.56</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0042110:8-cell proliferation</td>
<td>2</td>
<td>0.02273043</td>
<td>83.88</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0042392:9-homeostatic process</td>
<td>5</td>
<td>0.02380206</td>
<td>4.31</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0050778:positive regulation of immune response</td>
<td>3</td>
<td>0.02762708</td>
<td>11.10</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0044949:regulation of transcription</td>
<td>9</td>
<td>0.05079453</td>
<td>2.03</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0007160:8-transmembrane receptor protein tyrosine kinase signaling pathway</td>
<td>3</td>
<td>0.05171191</td>
<td>7.86</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0030183:8-cell differentiation</td>
<td>2</td>
<td>0.0846756</td>
<td>21.88</td>
</tr>
</tbody>
</table>

#### G2 TN (compared with G1 and G3 TN)

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Count</th>
<th>PValue</th>
<th>Fold Enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0002495:8-antigen processing and presentation of peptide antigen via MHC class II</td>
<td>4</td>
<td>8.59E-05</td>
<td>44.70</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0002504:8-antigen processing and presentation of polypeptide antigen via MHC class II</td>
<td>4</td>
<td>1.47E-04</td>
<td>37.64</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0030036:8-actin cytoskeleton organization</td>
<td>6</td>
<td>0.0216261</td>
<td>6.50</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0007159:8-leukocyte adhesion</td>
<td>3</td>
<td>0.0262056</td>
<td>38.31</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0008048:8-regulation of actin polymerization or depolymerization</td>
<td>4</td>
<td>0.0267285</td>
<td>14.30</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0030029:8-actin filament-based polymerization</td>
<td>6</td>
<td>0.0286076</td>
<td>6.10</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0012271:8-regulation of protein polymerization</td>
<td>4</td>
<td>0.0381628</td>
<td>12.55</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0043257:8-regulation of protein complex assembly</td>
<td>4</td>
<td>0.0552056</td>
<td>11.00</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0032535:8-regulation of cellular component size</td>
<td>5</td>
<td>0.0120336</td>
<td>5.55</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0021510:8-biological adhesion</td>
<td>9</td>
<td>0.0120698</td>
<td>2.86</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0019882:8-antigen processing and presentation</td>
<td>4</td>
<td>0.01228952</td>
<td>8.22</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0002573:8-myeloid leukocyte differentiation</td>
<td>3</td>
<td>0.01504575</td>
<td>15.78</td>
</tr>
</tbody>
</table>

#### G3 TN (compared with G1 TN)

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Count</th>
<th>PValue</th>
<th>Fold Enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0001878:8-response to yeast</td>
<td>3</td>
<td>1.46E-04</td>
<td>154.41</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0006920:8-response to fungus</td>
<td>3</td>
<td>0.00129085</td>
<td>54.50</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0008587:8-positive regulation of endocytosis</td>
<td>3</td>
<td>0.00514435</td>
<td>27.25</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0022508:8-positive regulation of transport</td>
<td>4</td>
<td>0.00833016</td>
<td>9.36</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0006959:8-phagocytosis</td>
<td>3</td>
<td>0.0104953</td>
<td>18.91</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0006952:8-immune response</td>
<td>6</td>
<td>0.01549443</td>
<td>3.93</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0006962:8-regulation of vesicle-mediated transport</td>
<td>3</td>
<td>0.0281563</td>
<td>11.16</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0016337:8-cell adhesion</td>
<td>4</td>
<td>0.03837022</td>
<td>5.23</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0032760:8-positive regulation of tumor necrosis factor production</td>
<td>2</td>
<td>0.0465453</td>
<td>41.18</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0009113:8-phagocytosis, engulfment</td>
<td>2</td>
<td>0.0465453</td>
<td>41.18</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0006952:8-defense response</td>
<td>5</td>
<td>0.05295938</td>
<td>3.45</td>
</tr>
</tbody>
</table>

#### G1 TN (compared with G2 and G3 TN)

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Count</th>
<th>PValue</th>
<th>Fold Enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0042110:8-T cell activation</td>
<td>17</td>
<td>1.33E-08</td>
<td>8.18</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0002520:8-immune system development</td>
<td>3</td>
<td>3.38E-08</td>
<td>3.72</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0048534:8-hematopoietic or lymphoid organ development</td>
<td>24</td>
<td>2.28E-07</td>
<td>3.60</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0030217:8-cell differentiation</td>
<td>12</td>
<td>1.57E-06</td>
<td>6.66</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0042113:8-cell activation</td>
<td>12</td>
<td>2.04E-06</td>
<td>6.49</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0046051:8-lymphocyte proliferation</td>
<td>8</td>
<td>5.10E-05</td>
<td>8.04</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0042110:8-cell proliferation</td>
<td>5</td>
<td>1.30E-04</td>
<td>17.58</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0050868:8-positive regulation of cell activation</td>
<td>31</td>
<td>1.35E-04</td>
<td>4.60</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0051249:8-regulation of lymphocyte activation</td>
<td>13</td>
<td>1.57E-04</td>
<td>3.81</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0005081:8-antigen receptor-mediated signaling pathway</td>
<td>7</td>
<td>1.78E-04</td>
<td>8.21</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0021510:8-V(D)J recombination</td>
<td>8</td>
<td>9.07E-04</td>
<td>5.32</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0050856:8-regulation of T cell receptor signaling pathway</td>
<td>7</td>
<td>0.0135294</td>
<td>5.68</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0030183:8-cell differentiation</td>
<td>4</td>
<td>0.00391788</td>
<td>12.06</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:003084:8-regulation of B cell activation</td>
<td>22</td>
<td>0.0092477</td>
<td>1.83</td>
</tr>
<tr>
<td>GOCHEM_BP_FAT</td>
<td>GO:0045160:8-cell fate commitment</td>
<td>4</td>
<td>0.0227069</td>
<td>6.49</td>
</tr>
</tbody>
</table>
regulated in both (Fig. 13E). Moreover, their pattern of expression is qualitatively linked to that of the red subset. Quantitatively, some genes are more highly expressed in the G1 Ly6D+ cells (Fig. 13D, upper panels) whereas others in the G2 Ly6D+ and G1 TN cells (Fig. 13D, lower panels). Among the latter ones, there are T-cell related genes such as Notch1, Lck, Rhoh, Ctlal2a, Ctlal2b, Gata3, Lat or Zap70 (Fig. 13D, lower-left panels), thereby indicating that the orange and purple cells might retain T-cell developmental potential. As a conclusion, these two groups have a lymphoid genetic profile that is not resolved into any particular lineage, presenting B and T enriched biological processes (Table 4D) and co-expressing B-, T- and lymphoid-genes (Il7r, Dntt or Lax1, Fig. 13D upper left). Finally, there are genes expressed in both green and blue cells (Fig. 13E). These are mostly myeloid related genes (Csfr1, Ccr2, Ifi30 or Ctsh) because although the blue cells are some way along the cDC lineage, as immature DC they have myeloid properties such as phagocytosis.

In summary, the single-cell transcriptomic analysis of the Ly6D+ and TN EPLM subpopulations reveals that: i) the clustered groups of cells have distinct genetic signatures (summarized in Fig. 14A) with a remarkable lympho-myeloid separation and ii) the degree of heterogeneity in the entire Ly6D+ and TN populations is reflected into their subgroups’ genetic profiles. Thus, both Ly6D+ subsets present a lymphoid genetic profile (with the G1 Ly6D+ cell in a more advanced developmental stage towards the B-cell lineage), whereas those of the TN present signatures to both the lymphoid (G1 TN) or myeloid (G2 and G3 TN) lineages, including some with a cDC lineage profile (G3 TN).

6.7 Expression of lymphoid and myeloid genes is mutually exclusive in single EPLM cells

We have observed that the same group of cells can co-express genes of different lymphoid or myeloid lineages. For instance, the orange G2 Ly6D+ cells express both the T-cell master regulator Notch1 and the early B-cell transcription factor Ebf1 (Fig. 13A,D), whereas the green G3 TN cells express both the neutrophil marker Elane and the macrophage colony-stimulating factor receptor Csf1r (Fig. 13C,D). In order to elucidate if the previous expression patterns also occur at the
single-cell level, we plotted the expression levels of representative pairs of transcripts, each characteristic of different lineages. We observed that a large
proportion of the $Ebf1^+$ cells also expressed $Notch1$ (75.6%, Fig. 14B left). The co-expression level was also high when comparing the immunoglobulin $\alpha$ chain of the pre-BCR complex ($Ig\alpha$ or $Cd79a$) with the T-cell tyrosine kinase $Lck$ (75% of the $CD79a^+$ and 28.5% of the $LcK^-$ cells, Fig. 14B right) Moreover, when examining neutrophil-monocyte/macrophage lineages, a high proportion of single-cells co-expressed $Elane$ and $Csfr1$ (76.9% of the $Elane^+$ and 23.3% of the $Csfr1^+$ cells, Fig. 14C). Therefore, the EPLM progenitor population contains single cells with mixed-lineage states within the lymphoid (B and T) and myeloid (granulocyte and monocyte/macrophage) lineages.

**Figure 15.** Marker candidates to isolate either the G1 TN (A) or the G1 Ly6D$^+$ group of cells (B). PCA plots generated as in Fig. 12A. Colour represents the expression level in log$_{2}$FPKM for each indicated gene per cell. Shape indicates the cell type: circles Ly6D$^+$, triangles TN.
In contrast, we detected distinct lymphoid and myeloid specification for the EPLM subgroups, with the cells on the left part of the PCA plot myeloid primed (G2 and G3 TN), whereas those in the centre (G1 TN and G2 Ly6D+) and on the right (G1 Ly6D-), lymphoid primed (Fig. 14A). This marked lympho-myeloid separation was confirmed at the single-cell level since we did not encounter significant co-expression of early myeloid (Cebpa and Ctsg) and lymphoid (Rag1 and Il7r) specification genes (Fig. 14D). The mutual exclusive expression of lymphoid and myeloid genes indicates that the EPLM might be composed of a mixture of cells with either lymphoid or myeloid priming and, that the multilineage developmental potential observed for the TN EPLM subpopulation is possibly not contained in the same single cell.

6.8 Selective markers for each subgroup. TdT the best candidate to separate cells with lymphoid from those with myeloid genetic profiles

After identifying selectively expressed markers for each of the clustered groups of cells (Fig. 13), we selected the best candidates in order to isolate the subgroups of interest and validate their genetic signatures with functional assays. We first isolated the G1 TN purple cells since we hypothesized that this TN fraction could be the precursor of the Ly6D+ EPLM subpopulation. For that, we performed and extensive screening using the PCA plot as a powerful way of interrogating the gene expression of any gene across the dataset and the best candidate was the Dntt gene (Fig. 15A upper left). In the PCA plot with cells coloured according to the Dntt transcript level (expressed in log$_2$FPKM), we observe that all G1 and G2 Ly6D+ cells (circles) are Dntt positive as well as the vast majority of G1 TN cells (purple in Fig. 13D upper left panel and Fig. 14A) with the exception of the small upper cluster of cells (box in the PCA plot in Fig. 15A). The Dntt gene encodes the Terminal deoxynucleotidyl Transferase (TdT) enzyme, a non-template polymerase that catalyzes the addition of random nucleotides at the junction of rearranged genes of both B- and T-cell receptors, thereby ensuring a highly diverse B- and T-cell receptor repertoire.
We checked the TdT protein expression by intracellular staining of the EPLM progenitor population and confirmed by flow cytometry that TdT can split the TN (B220⁺ CD117⁺⁺ CD19⁻ NK1.1⁻ Ly6D⁻ SiglecH⁻ CD11c⁻) cells into two almost equal fractions (50.5% TdT⁺) and that it is expressed by nearly all (87.8%) Ly6D⁺ EPLM (Fig. 16A). Therefore, TdT is potentially an excellent marker to isolate the TN TdT⁺ cells, equivalent to the G2 TN cluster identified in our bioinformatics analysis, and to test if their developmental potential matches that of their observed lymphoid genetic signature. Unfortunately, TdT is not a cell-surface marker and cannot be directly used to sort TdT positive cells by FACS. Moreover, there is no TdT reporter mouse available. For that reason, our laboratory is in the process of generating a TdT

Figure 16. Expression of TdT, CD115 and Ebf1 within the Ly6D⁺ and TN EPLM subpopulations. (A-C, upper panels) Representative FACS plot showing expression of TdT (A), CD115 (B) or Ebf1 (C) protein within the Ly6D⁺ and TN EPLM of Flt3Ltg (A,B) or Flt3Ltg-Ebf1⁻/⁺ (C). (A-C, lower panels) Percentages of TdT (A), CD115 (B) or Ebf1 (C) expressing Ly6D⁺ and TN EPLM from Flt3Ltg (n=3, A; n=4, B) or Flt3Ltg x Ebf1⁻/⁺ (n=3, C). Shown as mean±SEM. (D) Construct to generate the Ebf1⁻/⁺ reporter mice as reported in [112].
reporter mouse, which will shed light to this and other ongoing projects in the near future.

As an alternative to TdT, we screened for cell surface candidates that would enable us to enrich for the G1 TN fraction of cells. The most promising were those encoded by Csf1r, Ly6c2 and Ccr2 genes (Fig. 15A), whose expression is mostly overlapping with the blue G2 and green G3 TN cells (Fig. 13E and Fig. 14A). Moreover, as shown in the PCA plots (Fig. 15A) they are not expressed by exactly the same single cells, suggesting that they could be used in combination in order to exclude a major fraction of cells belonging to the G2 and G3 TN clusters. We attempted several protein surface stainings of the EPLM with individual or combinations of CD115 (or CSF-R, encoded by Csf1r gene), Ly6C2 and/or Ccr2 markers. However, we only detected significant expression of the CD115 cell surface protein, with about one third of the TN EPLM subpopulation being CD115+ (Fig. 16B).

We also wanted to isolate the G1 Ly6D+ red cells in order to validate their B-cell specification. For that, the two best candidates were the B-cell transcription factor Ebf1 and the BCR complex-associated protein α chain Cd79a (Fig. 15B and Fig. 13A). We had access to the Ebf1ihCd2/+ mice provided by Prof. Meinrad Busslinger and co-workers [111]. These mice had been generated by inserting an internal ribosome entry sequence (IRES)-hCd2 (ihCd2) reporter gene into the 3’ untranslated region of the Ebf1 gene (Fig. 16D). We next crossed the Ebf1ihCd2/+ reporter to the Flt3Ltg mice, thereby generating the Ebf1ihCd2/-Flt3Ltg mice. EPLM staining of the latter mice revealed that on average 57.5% of Ly6D+ cells express hCD2 and therefore Ebf1 (Fig. 16C).

6.9 Excluding CD115+ cells enriches for TN cells with lymphoid profile

In order to assess if the detected fraction of CD115- TN cells possesses the lymphoid genetic signature observed in the single-cell RNA sequencing experiment, we performed several assays. First, we examined the B-cell, T-cell and myeloid precursor frequencies of the two TN (CD115+ and CD115-) fractions. For that, we sorted them from Flt3Ltg mice and plated graded numbers of cells with the appropriate cocktail of stromal cells and cytokines (same experimental procedure
Figure 17. CD15 TN fraction contains the lymphoid profile. (A) Limiting dilution analysis of CD115\(^+\) TN and CD115\(^-\) TN for B-cell, T-cell or myeloid developmental potentials. EPLM subpopulations were sorted from Fli3Ltg (2 pooled mice) and plated at the indicated concentrations on either OP9 stromal cells together with IL-7, OP9-DL1 stromal cells in the presence of IL-7, or ST2 stromal cells. After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1 and ST2 cell cultures), B-cell, T-cell or myeloid clones were scored using an inverted microscope. One representative experiment is shown out of 2. (B) Heatmap of expression of lineage-specific genes as in Fig. 9B. The colour gradient illustrates the centred gene expression level (log\(_2\)CPM, mean=0 across all samples and per gene). (C) Reconstitution of B-cell compartment in sub-lethally irradiated B6 Rag2-deficient mice with 5\(\times\)10\(^3\) CD115 TN cells from WT mice. Left panel: CD19 versus IgM expression on spleen cells 3 weeks after transfer. Also CD21 and CD23 expression on gated CD19\(^+\) spleen cells are shown (middle panel). Right panel: Quantification of CD19\(^+\) IgM\(^+\) populations presented as frequency of live cells. CD115 TN (n=5). Shown as mean±SEM.
explained in results section 2 and Fig. 5). The results indicated that the CD115− TN fraction had lymphoid potential, poor B-cell precursor frequency (< 1 in 50) and strong T-cell precursor frequency (1 in 7), whereas the CD115+ TN did not (Fig. 17A upper and middle panel). Moreover, the latter population possessed considerably higher myeloid developmental potential (1 in 2 for CD115+ TN versus 1 in 20 for CD115− TN) (Fig. 17A lower panel). Although low, the myeloid potential observed in the CD115− TN could be explained because this fraction contains the G1 TN cells as well as others of the G2 and G3 clustered groups (Fig. 15A upper-right panel) that showed a myeloid genetic signature. We believe that using TdT as a marker we would be able to discriminate the cells with myeloid developmental potential.

The observed lymphoid developmental potential of the CD115− TN cells would appear to be in line with their genetic signature, as revealed by bulk RNA sequencing. Briefly, we compared the previously sequenced Ly6D+, TN and Pro-B populations (Fig. 9) with the newly identified CD115− TN fraction. In Fig. 17B, a heatmap generated with the same genes as was shown in Fig. 9B, indicates that the CD115− TN subset has a gene expression pattern more similar to the Ly6D+ than the entire TN population has to the Ly6D+ cells, especially for T-cell related genes. Finally, although the CD115− TN subset had a poor B-cell precursor frequency in vitro, they were able to reconstitute the splenic B-cell compartments of Rag2-deficient mice (Fig. 17C). This finding shows their in vivo B-cell developmental potential and suggests that there might be a factor that is necessary to support B lymphopoiesis missing in the in vitro cultures.

Collectively, these results indicate that the CD115 marker identified by single-cell RNA sequencing is able, as predicted by the genetic profiles of the clustered groups of cells (Fig. 14A), to exclude cells with myeloid precursor frequency and thereby enrich for those having lymphoid profiles and being the precursor of the Ly6D+ EPLM subpopulation.

6.10 Ebf1 enriches for Ly6D+ cells with B-cell potential

In order to elucidate if the newly identified Ebf1+ Ly6D+ subset is B-cell primed as predicted by the single-cell RNA sequencing experiment, we assessed their
Since the entire Ly6D\(^+\) EPLM subpopulation is already lymphoid restricted, we focused our analysis on examining in vitro and in vivo developmental potentials. Ebf1\(^+\) Ly6D\(^+\) fraction is enriched in B-cell whereas Ebf1\(^-\) Ly6D\(^+\) fraction in T-cell developmental potential. EPLM subpopulations were sorted from Flt3Ltg-Ebf1\(^{ihCd2/+}\) (2-pooled mice) and plated at the indicated concentrations on either OP9 stromal cells together with IL-7 or OP9-DL1 stromal cells in the presence of IL-7. After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1), B-cell or T-cell clones were scored using an inverted microscope. One representative experiment is shown out of 2.
the B-cell and T-cell precursor frequencies of the two (Ebf1⁺ and Ebf1⁻) Ly6D⁺ fractions by limiting dilution (Fig. 18A). The results indicate that although both subsets have *in vitro* B and T cell developmental potential, the Ebf1⁺ Ly6D⁺ fraction has a higher B-cell precursor frequency (1 in 5 for Ebf1⁺ Ly6D⁺ versus 1 in 21 for Ebf1⁻ Ly6D⁺), whereas the Ebf1⁻ Ly6D⁺ fraction has a higher T-cell precursor frequency (almost 1 in 2 for Ebf1⁻ Ly6D⁺ versus 1 in 11 for Ebf1⁺ Ly6D⁺) (Fig. 18A). These results were confirmed by their differential abilities to reconstitute mouse B- and T-cell compartments *in vivo* (Fig. 18B,C). Whereas the Ebf1⁺ Ly6D⁺ subset had all the B cell reconstitution potential (2/4 vs 0/5), the Ebf1⁻ Ly6D⁺ were the only cells able to reconstitute the thymus with TCRβ⁺ T cells (2/5 vs 0/4) (Fig. 18C). Overall, the genetic profiles of the two Ly6D⁺ clusters (red and orange in Fig. 14A) predicted by single-cell RNA sequencing, and exemplified with the Ebf1 marker, are reflected in their developmental potentials.
7. Discussion

The identification of two complementary clonogenic hematopoietic progenitors, CLP in 1997 [27] and CMP in 2000 [28], with the capacity to differentiate into all cells of the lymphoid and myeloid lineages respectively, resulted in the so-called “classical” model of haematopoiesis found in most current textbooks. This classical model reflects a hierarchical decision-making process whereby early multipotent progenitors make an early and irrevocable decision to differentiate towards either the lymphoid or the myeloid lineages. Since then, the classical model of haematopoiesis is in constant revision as a consequence of finding numerous additional progenitor cell types and alternative routes of differentiation. For instance, the identification of the LMPP [35] suggests an earlier branching of the megakaryocyte/erythrocyte pathway to the lympho-myeloid lineages, before the branching of the two latter lineages, and is supported by several later studies [70, 135]. Moreover, the irreversible and strict compartmentalization of haematopoietic progenitor cells was first challenged by the finding that in WT mice, committed precursor B cells regain multipotentiality upon conditional inactivation of the Pax5 gene [31]. This suggested that there might be a greater degree of developmental plasticity in haematopoiesis than previously thought. Subsequently, our laboratory detected in the BM of wild type mice a cell with similar properties to the Pax5−/− Pro-B cells that we called “Early Progenitor with Lymphoid and Myeloid potential” (EPLM) [33]. This was an uncommitted and multipotent haematopoietic B220+ CD117int CD19− NK1.1− progenitor with combined lymphoid and myeloid developmental potential, therefore challenging the lympho-myeloid dichotomy supported by the classical model of haematopoiesis.

In the present study, we have interrogated the multipotentiality of the EPLM in detail and assessed its potential heterogeneity. First, we found that EPLM expressed heterogeneous levels of three cell surface markers Ly6D, SiglecH and CD11c, resulting in subdivision of EPLM into at least four subpopulations (Fig. 4D). Moreover, this EPLM subdivision was still present in Flt3Ltg mice, in which the numbers of all EPLM subpopulations were significantly increased almost equally (about two orders of magnitude, Fig. 4E), thus making the Flt3Ltg mouse a good
model to further study EPLM subpopulations. When we assessed the in vitro and in vivo developmental potential of individual subsets by limiting dilution and reconstitution assays respectively, we observed that EPLM heterogeneity was reflected in different sets of potentials. Briefly, SiglecH+ and CD11c+ subpopulations could not generate lymphoid cells and, as their cell surface marker profile suggests, they could be (at least a fraction) already committed to the pDC and cDC lineages. Ly6D+ cells were lymphoid restricted and TN showed trilineage (B, T and myeloid) developmental potential (Fig. 5) although with lower B- and T-cell precursor frequencies compared with their Ly6D+ counterparts (Fig. 5A,B). Of note, Ly6D+ and TN cells from Flt3Ltg mice showed lower B-cell potential but higher T-cell precursor frequencies (Fig. 5A,B), as well as increased T-cell reconstitution efficacy in the thymus (Fig. 6D). One explanation for this finding could be the lower levels of Pax5 and Ebf1 B-cell transcription factors detected in Ly6D+ cells from Flt3Ltg mice compared to their WT counterparts (von Muenchow et al., Appendix paper 1). However, Flt3L is not an instructive factor for B-cell commitment of these progenitors. Instead, the main effect of Flt3L is the induction of their proliferation (von Muenchow et al., Appendix paper 1). Therefore, it could well be that an increased percentage of cycling progenitors results in a decreased fraction initiating the B-cell developmental program, thus explaining the proportional reduction of Ebf1- and Pax5-expressing Ly6D+ cells and the increased ability of these progenitors to give rise to T cells. However, it is important to mention that, in Flt3Ltg mice, the total number of B-cell progenitors is not affected due to the general expansion of the EPLM compartment (Fig. 4). Since Flt3L is reported to be important for pDC development [109, 136, 137] and the EPLM subpopulations might have pDC developmental potential, we hypothesize that pDC progenitors might be expanded in the transgenic mice, thereby providing an alternative explanation for the reduced B-cell potential. Although a clonal assay to assess quantitative pDC precursor frequency does not yet exist, it would be of interest to investigate the pDC developmental potential of EPLM subpopulations from WT and Flt3Ltg mice.

Another striking observation from the developmental potential assays was that, in spite of the poor in vitro capacity of TN cells to generate B cells (Fig. 5A), they were nevertheless able to reconstitute the splenic B-cell compartments of all
injected mice (Fig. 6C). This result therefore suggests that in in vitro cultures, there might be a factor missing that is necessary to support B lymphopoiesis. The opposite was observed for the T-cell potential in that this was more efficient in vitro than in vivo (Fig. 5B and Fig. 6D). A key difference in the two experimental systems used to determine T-cell potential is the availability of Notch ligands. Whereas the in vitro system generates highly sensitive readouts that take advantage of the T-cell lineage-promoting activity of Notch1 signalling via culturing progenitors on feeder cells expressing Notch ligands [138], the intravenous injection system depends on progenitors travelling to the thymus prior to receiving Notch ligand signalling upon thymus entry. Experimentally, this limitation could be overcome if progenitors were directly injected intrathymically. However, EPLM are progenitor cells residing in the bone marrow and therefore it is probable that although Ly6D+ cells show robust in vitro T-cell potential, their “default” physiological fate, when no other lineage commitment signal, such as Notch ligand, interferes, might be to become a B cell. Finally, we observed similar in vivo capacity of Ly6D+ and TN progenitors to reconstitute the splenic B-cell compartment (Fig. 6C). To further confirm this, it would be worthwhile injecting limiting numbers of progenitors and monitoring their reconstitution at different time points. Overall, these results and associated discussions raise the fundamental question of which assays truly represent the physiological potential of progenitor cells.

Transcriptomic analysis of the Ly6D+ and TN EPLM subpopulations showed that their distinct developmental potentials were supported by differential gene expression programs. Thus Ly6D+ EPLM had a lymphoid molecular priming and transcriptomic profile closer to committed Pro-B cells, whereas the TN had a more myeloid genetic signature (Fig. 9 and Table 3). Even though TN cells up-regulate myeloid genes involved in innate biological processes, we were able to demonstrate the precursor-product relationship between TN and Ly6D+ EPLM subpopulations, with the TN (or at least a fraction thereof) being the precursor of the Ly6D+ subpopulation, and therefore explaining the slower kinetics and decreased efficiency of the latter progenitor differentiating into CD19+ B cells (Fig. 13). Collectively, the previous data reveals that the TN EPLM subpopulation is a lympho-myeloid
progenitor, whereas the Ly6D\textsuperscript{+} EPLM subpopulation is likely to be the direct precursor of CD19\textsuperscript{+} committed Pro-B cells.

The emergence of high throughput methods enabling the investigation of whole-genome or whole-transcriptome profiles from single cells, boosts the existing debate regarding the heterogeneity of apparent phenotypically homogeneous progenitor cells that have multiple lineage potentials. To shed light onto this active debate, by performing single-cell RNA sequencing we provide herein a detailed analysis of the heterogeneity of the Ly6D\textsuperscript{+} and TN EPLM subpopulations. This study captured 365 single-cell gene expression snapshots of the Ly6D\textsuperscript{+} (152) and TN (213) transcriptional landscapes. Principal component analysis is a good method to visualize and interpret high dimensional data (e.g. 365 single cells x 14814 detected genes). This method maximizes the variance of the features (genes in our case). Therefore, in order to obtain informative components (or dimensions) in the plots, it is common practise to select a subset of the original gene set that contains most of the variation among the samples [53, 139, 140]. One strategy is to perform an unsupervised selection of, for example, the 50\% of genes with the highest variance across the analysed dataset, as was done for the visualization of the bulk RNA sequencing experiment. However, to analyse the single-cell data we decided to concentrate on the 1008 genes that we identified as differentially expressed between the Ly6D\textsuperscript{+} and TN populations. Our reasoning was that this set of genes would provide insights into the intra-population variation. This is because the population approach obscures molecular heterogeneity by averaging gene expression. By analysis of bulk populations, high expression of a particular gene in an individual subset will be interpreted as general expression by all cells. As suggested by their branching structure in the PCA and their low cell-to-cell transcriptome correlation our initial analysis indicated that the TN subset is a more heterogeneous population than their Ly6D\textsuperscript{+} partners (Fig. 11). Moreover, the PAM clustering based on the selected subset of genes, partitioned the Ly6D\textsuperscript{+} into two major clusters whereas the TN was subdivided into three robust groups. Of note, Ly6D\textsuperscript{+} and TN cells had not been completely segregated in the PCA (Fig. 11B), thereby suggesting that the overlapping cells might be in a related cellular state. This was confirmed by the analysis of the two clustered groups corresponding to the convergent area, namely
G2 Ly6D⁺ and G1 TN (Fig. 12A), revealing that they have the highest transcriptome correlation among the subgroups (0.721, Fig. 12B) and the lowest number of DEG (25, Fig. 12C). Therefore, the G1 TN is likely to be the fraction that we previously observed experimentally to be the precursor of Ly6D⁺ cells. However, our latter in silico analysis performed at the whole-transcriptome scale of the G2 Ly6D⁺ and G1 TN subgroups, which are sorted as two phenotypically distinct populations, suggests that they should be considered as one subset. This finding highlights the limitations of accurately defining complete cellular identities by relying on expression of few cell surface markers, the so-called “top-down” approach [141]. This notion has been manifested in other studies such as when Paul et al. suggested that the standard gating for sorting MEP might be better termed erythrocyte progenitor (EP) gating [69].

When performing cell clustering, it is of pivotal importance to investigate the nature of the subgroups in order to determine if the cells are clustered as a result of stochastic fluctuations of the transcriptome and certain cellular states (e.g. cell cycle status or cellular stress), or instead, the clustering reflects gene expression variation that might be significant for the biological function being tested [142]. In our case, the hypothesis is to ask whether individual EPLM are multipotent or whether as a population they are composed of a mixture of cells with distinct lineage developmental potentials. Therefore, we sought to determine if the clustered groups of cells reflect genetic signatures of progenitors from distinct haematopoietic lineages. For that, we extended our analysis to all detected genes (14,814) across the 365 single-cells and examined the DEG and enriched biological processes defining each subgroup (Table 4 and supplementary excel file 3_DEGlists_scRNAseq). With this approach, we were able to unravel marked genetic biases between EPLM indicative of molecular priming towards distinct fates. Whereas the G1 Ly6D⁺ subgroup showed a strong B-cell genetic signature (red Fig. 14A) with robust expression of B-cell related genes characteristic of the Pro-B cell stage (Cd79a, Vpreb genes, IgIl1, Cd19, Ebf1 or Blnk Fig. 13A) and B-cell enriched biological processes (Table 4A), the G2 TN subgroup, which has the most distant transcriptome, revealed a cDC genetic signature (blue Fig. 14A) with consistent expression of genes (H2-genes, Cd74, Ciita, Id2 or Batf3 Fig. 13B) and enriched biological processes (Table 4B).
either related with antigen processing and presentation or necessary for cDC development. The G3 TN subgroup exhibited a myeloid genetic signature (green Fig. 14A), up-regulating myeloid genes (Mpo, Ctsg, Prtn3, Elane, Cx3cr1, Cebpa or Csfr1 Fig. 13C,E) related with innate processes (Table 4C). Finally, the G2 Ly6D⁺/ G1 TN subset (orange and purple respectively in Fig. 14A) showed a lymphoid genetic signature (Fig. 13D and Table 4D) with a B-cell specification milder to that of the G1 Ly6D⁺. Taken together, these results indicate that, as reflected by their central location in the PCA, these cells might me in an intermediate state while exhibiting promiscuous expression of B- and T-cell genes. Therefore, the previous data suggest functional heterogeneity among the EPLM subpopulations and validates single-cell RNA sequencing as a powerful technology to dissect molecular heterogeneity of a previously considered homogeneous population via generating biologically meaningful clusters of cells. In addition, numerous studies have reported the use of this technology with similar purposes [46, 53, 57, 67-69, 139, 143-145]. For instance, Gren et al. reveal transcriptomic variation within the traditionally classified classical, intermediate and non-classical monocytes [67]; Paul et al. report heterogeneity in myeloid progenitors [69]; Kowalczyk et al. find extensive transcriptome variability among HSCs [145]; and Drissen et al. subdivided the pre-GM population into a Gata1⁺ pre-GM fraction generating mast cells, eosinophils, megakaryocytes and erythroid cells and a Gata1⁻ pre-GM subset generating monocytes, neutrophils and lymphocytes [139]. These publications challenge the existence of a clonal myeloid progenitor generating all innate myeloid immune cell types.

We also investigated lineage priming at the single-cell level and found a significant proportion of single cells co-expressing early B- and T-cell (Fig. 14B) or granulocyte and monocyte/macrophage (Fig. 14C) specification genes. This is consistent with other studies where mixed lymphoid [126, 146] or myeloid [135, 147, 148] lineage patterns of gene expression are reported in single cells. However, whereas heterogeneity is well studied in myeloid progenitors, we are not aware of other reports addressing mixed lymphoid priming at the single-cell level and whole-transcriptome scale, thus highlighting the importance of our study. Strikingly, and in agreement with other reports [144, 149-151], we did not observe single cells with mixed myeloid and lymphoid lineage gene expression profiles (Fig. 14D).
Hans-Reimer Rodewald and co-workers made use of an Il7r fate mapping mouse line to determine the almost exclusive lymphoid progeny of Il7r expressing cells, thus further supporting our findings [149]. Nevertheless, we cannot exclude that, due to the “snapshot” nature of the transcriptomic analysis, as well as the medium throughput of cells analysed and the medium sequencing depth, we are missing a transient and presumably rare intermediate state with promiscuous lympho-myeloid gene expression. Previous data suggested that bifurcation of the lymphoid and myeloid molecular priming occurs before the EPLM stage. Therefore, the common or separate origin of the Ly6D+/G1TN (lymphoid primed) versus the G2/G3TN (myeloid primed) EPLM fractions is of interest and requires further investigation. In line with our findings, there is an increasing body of evidence supporting the notion that priming occurs much earlier in development than previously thought. Indeed, expression of lineage-affiliated genes has been reported as early as in the HSC stage, with various analyses indicating biases at the apex of haematopoiesis [135, 144, 152-158], thereby suggesting that the cellular output of the multipotent early progenitors is small [135].

One key aspect of single-cell RNA sequencing studies is the identification of markers representative of the identified clusters of cells in order to prospectively isolate them and perform functional assays to validate the predicted genetic signatures. In our study, we found that the reclassification of EPLM subpopulations was exemplified by expression of Dntt (encoding TdT), which was expressed in the Ly6D+ as well as the vast majority of G1 TN but not in G2 and G3 TN cells (Fig. 15A). We believe that TdT reporter mice (being currently generated in our laboratory), will provide the exciting opportunity not only to isolate the lymphoid-primed EPLM fraction but also to test and demonstrate, as previously suggested [159], if TdT+ CD19- cells constitute the earliest lymphoid progenitor in adult mice. Moreover, this reporter will be of importance in studies comparing fetal versus adult haematopoietic progenitors and to establish when TdT is first expressed in ontogeny.

Profiting the power of single-cell RNA sequencing to identify markers in an unbiased manner, we performed an extensive screening and found potential candidates (Fig. 15A,B). As a proof-of-concept, we made use of Ebf11hCd42/+ reporter mice in order to test the developmental potential of Ebf1+ EPLM (comprising most of
the G1 and some G2 Ly6D⁺ cells). As expected, the Ebf1⁺ Ly6D⁺ fraction exhibited higher B-cell precursor frequency and in vivo reconstitution potential than the Ebf1⁻ Ly6D⁺ fraction (Fig. 18). This result demonstrates that we can find markers from our single-cell transcriptomic analysis that, when tested, successfully predict developmental outcome. Although the G1 Ly6D⁺ fraction has a strong B-cell genetic specification and developmental potential, the observed marginal in vitro T-cell potential suggests that they are still not restricted to the B-cell lineage. This supports the concept that genetic specification precedes commitment [147, 160, 161] and that Ebf1 initiates B-cell specification whereas Pax5 acts later to drive commitment [126, 160]. Of note, in our single-cell analysis, we barely detected Pax5 expression (n=8), which could be due to the insufficient sequencing depth to detect low expressing transcripts, or the use of Flt3Ltg mice (in which Pax5 is lower at the EPLM stage, von Muenchow et al. Appendix paper 1). However, even though EPLM are sorted as CD19⁻, we consistently detected Cd19 transcript (n=19). Therefore, we propose the herein newly identified G1 Ly6D⁺ subset as the direct precursor of the first B-cell committed stage, namely the CD117⁺ CD19⁺ Pro-B cells.

Additionally, we tested the CD115 cell surface marker (encoded by Csf1r gene) identified in our single-cell RNA sequencing experiment and expressed by a good proportion of the G2 and G3 TN cells. Interestingly, our molecular and functional analysis revealed that compared with the entire TN population, CD115⁻ TN had reduced myeloid but enriched lymphoid properties (Fig. 17), thus confirming that priming at the molecular level correlates with the immediate differentiation bias of EPLM cells. Therefore, our data support intrinsic differentiation potentials through differential gene expression programs, although we do not exclude the contribution of other factors such as epigenetic regulation or extrinsic cytokine signalling.

Single-cell transcriptome profiling has already enabled advances to be made in a broad variety of fields. However, it implies a subsequent prospective strategy which, although useful for this and other research studies [65, 139], depends on the identification of cell-surface markers or the existence of reporter mice in order to translate the single-cell transcriptome findings. Additional genetic tools and integration of different omics analysis will certainly be developed, thus enhancing the contribution of single-cell analysis technology to the field of cell lineage
determination. Indeed, the first genome and transcriptome sequenced from the same individual cell already exists [162], enabling a direct comparison of genomic variation and transcriptome heterogeneity. The challenge will be to generate methods for RNA and protein measurements without destroying the cell of interest in order to relate gene expression with cell behaviour. These measurements will require transcriptional and translational reporters combined in single-molecule imaging and quantification experiments.

As a summary, in this thesis, by studying the previously reported multipotent and uncommitted B220^+ CD117^{int} CD19^- NK1.1^- EPLM progenitor, we have been able to: first divide the EPLM into four subpopulations with distinct lineage biases using the “top down” strategy. This is based on the expression of cell surface markers (Ly6D, SiglecH and CD11c in our case). Subsequently, we further investigated the EPLM heterogeneity using the alternative and complementary “bottom up” approach [141], in which high-dimensional molecular profiles measured for each single cell are used to classify distinct states and types. The latter strategy has revealed that: i) Ly6D^+ and TN EPLM subpopulations have distinct degrees of molecular and functional heterogeneity ii) the Ly6D^+ subset is composed of two lymphoid specified subgroups with a B-cell priming gradient iii) the TN subset is composed of three groups of cells with lymphoid or myeloid genetic signatures, including some cells with a cDC lineage profile iv) genetic signatures reveal a remarkable lympho-myeloid separation and v) this separation is better exemplified by Dntt instead of Ly6D, thereby redefining the initial “top down” classification. Taken together, these results indicate that the EPLM is a good example supporting the finding that previously characterized multipotent progenitor populations are in fact composed of mixtures of cells with differently restricted differentiation capacities. EPLM share their B220, CD117 and CD19 phenotype with Pax5^{-/-} Pro-B cell clones that exhibit multilineage differentiation capacities. However, despite this phenotypic similarity, we can conclude from the analysis of the developmental potential of individual subsets that EPLM are in fact distinct from Pax5^{-/-} Pro-B cells [30, 31, 33]. More generally, this study illustrates how single-cell expression profiling can be used to identify novel subpopulations and cellular intermediates that can then be prospectively isolated and characterized in functional assays.
Finally, in the context of constant and necessary revisions of the haematopoietic schemes, we argue in favour of a revised model of haematopoiesis in which lineage priming is initiated as early as at the HSC level and progressively increases with development, ultimately giving rise to the differentiated blood cell types. Thus, unlike the classical hierarchical and branching model of haematopoiesis, we support alternative representations like the pair-wise model ([Fig. 2; [41]] where the route a progenitor cell will take to arrive at a particular destination is more flexible than currently represented. This more flexible model better reflects haematopoietic versatility and can potentially accommodate the increasingly reported alternative differentiation pathways [36, 70, 139, 158]. However, our findings suggest a revision of the CLP/EPLM continuum arch of the pair-wise model. Now we know that the lymphoid priming of the EPLM resides in the TdT+ fraction (herein identified as the Ly6D+ as well as the G1 TN EPLM subpopulations), and therefore this is the fraction of EPLM that should be placed directly downstream of CLP rather than the entire EPLM population [33]. The location of the other EPLM fractions in the above scheme has yet to be defined and requires further investigation. Moreover, haematopoiesis seems to occur through a process of “graded commitment” [69, 135] and this is consistent with our finding of an increasing degree of B-cell priming within EPLM subsets (G2 Ly6D+/ G1 TN < G1 Ly6D+). How to illustrate the global haematopoietic process of “graded commitment” in the revised models of haematopoiesis remains an exciting challenge for the near future. Ultimately, this thesis makes a significant contribution to the study of the phenotypic and genotypic heterogeneity and lineage priming of progenitors during early stages of haematopoietic development.
8. References


9. Appendix

Further projects and publications in which I have participated

9.1 Permissive roles of cytokines Interleukin-7 and Flt3-ligand in mouse B cell lineage commitment

Lilly von Muenchow\(^1\), Llucia Albertí-Servera\(^1\), Fabian Klein\(^1\), Giuseppina Capoferri\(^1\), Daniela Finke\(^2\), Rhodri Ceredig\(^3\), Antonius G. Rolink\(^1\) and Panagiotis Tsapogas\(^1\)

\(^1\) Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland
\(^2\) Developmental Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland
\(^3\) Department of Biosciences, University of Galway, Galway, Ireland

Second author – In press (PNAS)

9.2 Reconstitution of a functional B-cell compartment in immunodeficient mice with pro-B cells propagated with or without stromal cells

Lilly von Muenchow\(^1\), Panagiotis Tsapogas\(^1\), Llucia Albertí-Servera\(^1\), Giuseppina Capoferri\(^1\), Marianne Doelz\(^1\), Hannie Rolink\(^1\), Nabil Bosco\(^{1,2}\), Rhodri Ceredig\(^3\) and Antonius G. Rolink\(^1\)

\(^1\) Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland
\(^2\) Present address: Nestlé Research Center, Lausanne, Switzerland
\(^3\) Department of Biosciences, University of Galway, Galway, Ireland

Third author – In press (EJI)
9.3 Versatility of stem and progenitor cells and the instructive actions of cytokines shape haematopoiesis

Geoffrey Brown¹, Ciaran J Mooney¹, Llucia Albertí-Servera², Lilly von Münchow², Kai-Michael Toellner¹, Rhodri Ceredig³ and Antonius G Rolink²

¹ School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
² Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland
³ Regenerative Medicine Institute, National Centre for Biomedical Engineering Science, Department of Physiology, National University of Ireland, Galway, Ireland

Third author - Published in Critical Reviews in Clinical Laboratory Sciences, July 2015
9.1 Paper 1:

Permissive roles of cytokines Interleukin-7 and Flt3-ligand in mouse B cell lineage commitment

Lilly von Muenchow\textsuperscript{1}, Lucia Alberti-Servera\textsuperscript{1}, Fabian Klein\textsuperscript{1}, Giuseppina Capoferri\textsuperscript{1}, Daniela Finke\textsuperscript{2}, Rhodri Ceredig\textsuperscript{3}, Antonius Rolink\textsuperscript{1} and Panagiotis Tsapogas\textsuperscript{1*}

\textsuperscript{1} Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Mattenstrasse 28, 4054, Basel, Switzerland;
\textsuperscript{2} University of Basel Children's Hospital, Basel, Switzerland; Developmental Immunology, Department of Biomedicine, University of Basel, Mattenstrasse 28, 4054, Basel, Switzerland;
\textsuperscript{3} National Centre for Biomedical Engineering Science, University of Galway, Galway, Ireland

* Corresponding author contact: Panagiotis Tsapogas panagiotis.tsapogas@unibas.ch

Acknowledgement
A.G.R. is holder of the chair in immunology endowed by L. Hoffmann – La Roche Ltd, Basel. This study was supported by the Swiss National Science Foundation and by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013 under Research Executive Agency grant agreement number 315902. R.C. was supported by Science Foundation Ireland under grant numbers SFI09/SRC/B1794 and SFI07/SK/B1233b.

Conflict of interest
The authors declare no commercial or financial conflict of interest.
ABSTRACT

Hematopoietic cells are continuously generated throughout life from hematopoietic stem cells, thus making hematopoiesis a favorable system to study developmental cell lineage commitment. The main factors incorporating environmental signals to developing hematopoietic cells are cytokines, which regulate commitment of hematopoietic progenitors to the different blood lineages by acting either in an instructive or a permissive manner. Flt3-ligand (FL) and Interleukin-7 (IL7) are cytokines pivotal for B cell development, as manifested by the severely compromised B cell development in their absence. However, their precise role in regulating B cell commitment has been the subject of debate. In the present study we assessed the rescue of B cell commitment in mice lacking IL7 but simultaneously over-expressing FL. Results obtained demonstrate that FL over-expression in IL7 deficient mice rescues B cell commitment, resulting in significant Ebf1 and Pax5 expression in Ly6D^+CD135^+CD127^-CD19^- precursors and subsequent generation of normal numbers of CD19^+ B cell progenitors, therefore indicating that IL7 can be dispensable for commitment to the B cell lineage. Further analysis of Ly6D^+CD135^+CD127^-CD19^- progenitors in IL7- or FL-deficient mice over-expressing Bcl2, as well as in IL7-transgenic mice suggests that both FL and IL7 regulate B cell commitment in a permissive manner; FL by inducing proliferation of CD135^-CD127^-Ly6D^-CD19^- progenitors and IL7 by providing survival signals to these progenitors.
INTRODUCTION

Hematopoiesis, the generation of all blood cells from hematopoietic stem cells (HSC), takes place continuously in the adult bone marrow. Accumulating evidence suggests that HSC generate the different hematopoietic lineages via oligopotent progenitors having limited self-renewal capacity and restricted developmental potentials. Activation of lineage-specific gene transcription in these progenitors eventually leads to their commitment to a particular lineage. Cytokines are the most prominent environmental factors regulating hematopoietic lineage commitment, doing so by acting either in an instructive or a permissive manner (1). In the instructive model, cytokines induce a signaling cascade in progenitors leading to the initiation of a lineage-specific gene program, typically through up-regulation and/or activation of transcription factors, eventually resulting in commitment to a particular lineage. In contrast, the permissive model advocates that commitment of progenitors to different lineages occurs in a cell-autonomous, stochastic manner, with cytokines acting as a selection rather than a commitment factor, promoting the survival and/or proliferation of a specific lineage at the expense of other lineages originating from the same progenitor. Elucidating the precise mode of action of cytokines is technically challenging and therefore the instructive versus permissive roles of cytokines is hotly debated (2-4). Although the permissive model was favored in the past, recent data provide solid evidence for the instructive action of several cytokines including M-CSF, G-CSF, EPO and Flt3-ligand (5-8). However, our understanding of how cytokines regulate hematopoiesis remains elusive, as different cytokines can act in various ways and their function might be cell-context dependent (9). Moreover, most studies to date have addressed cytokine regulated myeloid differentiation with relatively little information on lymphoid lineage commitment.

That Interleukin-7 (IL7) is a crucial cytokine for B cell generation is demonstrated by the dramatic defect in B cell generation in mice lacking either the cytokine (10) or its receptor (11). Interestingly, while human B cell progenitors are also responsive to IL7 (12), disruption of IL7 signaling caused by mutations does not ablate B cell development in man (13, 14). IL7 was initially identified as a growth factor for B cell progenitors (15) and early studies demonstrated that in vivo over-
expression of the pro-survival gene \( \text{Bcl2} \) did not rescue B cell development in the absence of IL7 signaling suggesting that IL7 acts in an instructive manner in B cell commitment (16, 17). The subsequent findings that uncommitted Common Lymphoid Progenitors (CLP) from \( \text{Il7}^{-/-} \) mice lacked expression of the transcription factor Early B-cell Factor 1 (Ebf1) (18) and that Ebf1 over-expression partially restored B cell generation from these CLP (19), led to the hypothesis that IL7, through Stat5 activation, instructs commitment to the B cell lineage by initiating Ebf1 expression in uncommitted progenitors. Supporting this hypothesis, a putative Stat5 binding site was later identified in one of the \( \text{Ebf1} \) promoters (20). However, a more recent study has shown that \( \text{Bcl2} \) can rescue B cell generation in a \( \text{Stat5} \) conditional knock-out mouse (21). Furthermore, the Ebf1-expressing fraction of CLP (Ly6D\(^+\) CLP) is dramatically reduced in \( \text{Il7}^{-/-} \) mice (22), therefore providing an alternative possibility for the reduced Ebf1 expression observed in \( \text{Il7}^{-/-} \) CLP. Interestingly, B cell lineage commitment is initiated at the molecular level in Ly6D\(^-\)CD19\(^-\) progenitors (23). Hence, while the importance of IL7 as a growth factor for committed B cell progenitors has been well established, it remains unclear whether it instructs oligo-potent progenitors to commit to the B cell lineage through \( \text{Ebf1} \) and \( \text{Pax5} \) up-regulation.

Ft\(3\)-ligand (FL), the only known ligand for the Flt3 receptor (CD135), is a cytokine important for the generation of many hematopoietic lineages and its function has gained much attention as mutations in FL signaling are commonly found in Acute Myeloid Leukemias (AML) (24). Committed B cell progenitors do not express CD135 since expression of the B cell commitment factor \( \text{Pax5} \) leads to its down-regulation (25). However, upon transplantation, bone marrow progenitors from \( \text{Flt3}^{-/-} \) and \( \text{Flt3l}^{-/-} \) mice reconstitute the B cell compartment poorly (26, 27), and FL was found to be essential for maintaining normal numbers of uncommitted B cell progenitors (28).

Recently, we described a FL-transgenic mouse model (hereafter \( \text{Flt3Ltg} \)) expressing high levels of FL \textit{in vivo}, which has enabled us to suggest an instructive role for FL in early stages of hematopoiesis (8). By breeding these mice with \( \text{Il7}^{-/-} \) mice we herein show that increased FL levels can rescue B cell commitment in CD135\(^-\)CD127\(^-\)CD19\(^-\) progenitors and restore early CD19\(^+\) B cell progenitor numbers.
in the absence of IL7 signaling, suggesting a permissive role for IL7 in B cell commitment. Further analyses of a combination of mouse genotypes over-expressing or lacking FL and IL7, as well as the pro-survival gene Bcl-2, have enabled us to identify a permissive role for both IL7 and FL in B cell commitment.
RESULTS

Increased \textit{in vivo} levels of FL rescue B cell commitment in \textit{Il7}\textsuperscript{-/-} Ly6D\textsuperscript{+}CD19\textsuperscript{-} progenitors

We have previously characterized an uncommitted B cell progenitor with combined lymphoid and myeloid potential (Early Progenitor with Lymphoid and Myeloid potential – EPLM) (29). EPLM can be further subdivided by SiglecH, CD11c, CD115 and Ly6D expression enabling us to identify the Ly6D\textsuperscript{+}SiglecH\textsuperscript{-}CD11c\textsuperscript{-}CD115\textsuperscript{-} fraction of EPLM (hereafter Ly6D\textsuperscript{+} EPLM) as the population containing most B cell potential, while being devoid of myeloid potential (Fig. S1B). This EPLM subpopulation is identified as Lin\textsuperscript{-}CD19\textsuperscript{CD117\textsuperscript{int}B220\textsuperscript{int}Ly6D\textsuperscript{CD135\textsuperscript{-}CD127\textsuperscript{+}} (Figure 1A) therefore partially overlapping phenotypically with Ly6D\textsuperscript{+} CLP (Fig. S1A) and pre-pro-B cells (30, 31). Ly6D\textsuperscript{+} EPLM numbers in \textit{Il7}\textsuperscript{-/-} and \textit{Flt3l}\textsuperscript{-/-} mice are significantly decreased compared to WT; 7-fold for \textit{Il7}\textsuperscript{-/-} and 13-fold for \textit{Flt3l}\textsuperscript{-/-} respectively and a similar dramatic decrease was observed in Ly6D\textsuperscript{+} CLP from both mutant mice (Figure 1B-C). FL deficiency also affected the numbers of Ly6D\textsuperscript{-} EPLM and CLP, while IL7 did not (Fig. S1C-D). Therefore, Ly6D\textsuperscript{+} EPLM/CLP represent the earliest developmental stage of the B cell pathway affected by the absence of IL7.

We have recently generated a mouse model expressing high \textit{in vivo} levels of FL (8). The progenitor compartment of these mice showed a dramatic increase in EPLM and CLP numbers, with their Ly6D\textsuperscript{+} fractions increased 90-fold and 28-fold respectively relative to WT (Figure 1D-E). We crossed \textit{Flt3Ltg} with \textit{Il7}\textsuperscript{-/-} mice to assess the extent to which increased FL levels could potentially rescue the loss of Ly6D\textsuperscript{+}CD19\textsuperscript{-} progenitors in \textit{Il7}\textsuperscript{-/-} mice. As shown in Figure 2 A-B, \textit{in vivo} over-expression of FL leads to a significant increase in \textit{Flt3Ltg-Il7\textsuperscript{-/-}} EPLM and CLP numbers, reaching levels of those in \textit{Flt3Ltg} mice. Crucially, a full rescue of Ly6D\textsuperscript{+} EPLM and CLP can be seen in these mice, with a striking 470-fold and 31-fold increase in numbers compared to their \textit{Il7\textsuperscript{-/-}} counterparts (Figure 2A-B). Furthermore, the numbers of the earliest committed CD19\textsuperscript{+}CD117\textsuperscript{-}IgM\textsuperscript{-} and CD19\textsuperscript{+}IgM\textsuperscript{-} B cell stages, since these cells require IL7 to
expand. As a consequence of this rescue in bone marrow B cell development, numbers of splenic marginal zone and follicular B cells were significantly increased in Flt3Ltg-II7−/− mice compared to II7−/− (Figure 2D). While thymic T cell development was not rescued in Flt3Ltg-II7−/− mice (Fig S3), a significant increase in splenic T cell numbers was observed (Fig. S4) as a result of their expansion upon FL overexpression (32).

To assess whether these rescued Flt3Ltg-II7−/− Ly6D⁺CD19⁺ progenitors could give rise to B cells in vitro, we sorted Flt3Ltg-II7−/− Ly6D⁺ EPLM and plated them at limiting dilution on OP9 stromal cells in the presence of IL7. As shown in Figure 3A, Flt3Ltg-II7−/− Ly6D⁺ EPLM could generate B cells at similar frequencies to their WT and Flt3Ltg counterparts, whereas the few II7−/− Ly6D⁺ EPLM isolated could not. A rescue in Ly6D⁺ EPLM was also observed when II7−/− mice were injected with FL (Figure 3E) and when plated under the same conditions these rescued Ly6D⁺ EPLM also showed a restored in vitro B cell potential (Fig. S5A). Further, when transplanted into irradiated Rag2−/− mice they were able to generate IgM⁺ B cells (Fig S5B-C). Thus, increased FL levels restore the generation of Ly6D⁺ progenitors, rather than merely expanding the few Ly6D⁺ EPLM/CLP found in II7−/− mice. Real-time quantitative PCR analysis of Ly6D⁺ EPLM from Flt3Ltg-II7−/− mice revealed significant expression of Ebf1, Pax5 and Foxo1 transcription factors’ mRNA in the absence of IL7 (Figure 3B). Ebf1 expression at the protein level was confirmed by intracellular FACS staining (Figure 3C-D). Even though the percentage of Flt3Ltg-II7−/− Ebf1⁺Ly6D⁺ EPLM did not reach WT levels, it was similar to the one found in Flt3Ltg mice, which produce IL7. Therefore, Ebf1/Pax5 expression and subsequent commitment to the B cell fate can occur in the absence of IL7 signaling, arguing against an instructive role of this cytokine in B cell commitment.

CD127 (IL7Rα) is a receptor shared between IL7 and thymic stromal lymphopoietin (TSLP), a cytokine capable of rescuing B cell development when over-expressed in the absence of IL7 (33). Since TSLP is produced by dendritic cells (34), which are dramatically expanded in Flt3Ltg mice (8), in vivo FL over-expression could lead to increased levels of TSLP thereby rescuing B cell development in Flt3Ltg-II7−/− mice. To investigate this possibility we injected II7−/− or II7rα−/− mice with FL as described above and assessed the rescue of Ly6D⁺ EPLM and downstream CD19⁺...
progenitors. FL injections into Il7−/− mice resulted in a significant increase in Ly6D+ EPLM and CD19+CD117+ B cell progenitors, comparable to the rescue observed in Flt3Ltg-Ill7−/− mice (Figure 3E). FL injected Il7ra−/− mice also demonstrated a significant rescue of Ly6D+ EPLM and CD19+CD117+ pro-B cells, indicating that the observed rescue of B cell commitment in Flt3Ltg-Ill7−/− mice is not mediated through the action of TSLP.

IL7 promotes survival, but not proliferation, of Ly6D+CD135+CD127−CD19− progenitors

Even though our Flt3Ltg-Ill7−/− mouse model suggests that IL7 is dispensable for B cell commitment, the dramatic decrease in Il7−/− Ly6D+ EPLM/CLP argues for a role of IL7 in the maintenance of this population when FL levels are limiting, by promoting either their survival or their proliferation. To investigate the potential role of IL7 as a survival factor for Ly6D+CD135+CD127−CD19− progenitors, we crossed Il7−/− mice with mice expressing the pro-survival gene Bcl2 (35). Bcl2tg-Ill7−/− mice demonstrated a minor but statistically significant 2.6-fold increase in Ly6D+ EPLM and 2.2-fold increase in Ly6D+ CLP numbers compared to Il7−/− mice (Figure 4A-B). Cell cycle stage analysis of Ly6D+ EPLM of these mice indicated that most of the cells rescued by Bcl2 are in a quiescent state (Fig. S6) and do not proliferate in response to cytokines, thereby compromising to some extent the rescue of these progenitors’ numbers. Importantly, when plated on OP9 stromal cells plus IL7, Bcl2tg-Ill7−/− Ly6D+ EPLM generated B cells at frequencies similar to WT mice (Figure 4C), indicating that these rescued Ly6D+ cells had B cell potential. Indeed, when analyzing bone marrow CD19+ committed progenitors, we could see a significant 68-fold increase in the earliest CD19+CD117+ pro-B cell compartment, compared to Il7−/− (Figure 4D). Due to the their quiescent state (Fig. S6 and (36)) and the IL7 dependence of their proliferation, Bcl2tg-Ill7−/− CD19+CD117+ numbers did not reach WT levels, whereas downstream CD19+ immature B cells showed a less pronounced, but significant rescue (Figure 4D). In the periphery of these mice, marginal zone and follicular B cell numbers were increased, whereas as previously reported (16), T cell numbers were rescued (Fig. S7). Therefore, providing an extra Bcl2-mediated survival signal in vivo partially rescues Il7−/− Ly6D+CD19− progenitors with B cell potential and restores
significantly the generation of CD19⁺ progenitors. This suggests a role for IL7 in facilitating the survival of Ly6D⁺CD135⁺CD127⁺CD19⁻ progenitors.

To evaluate the potential proliferative effect of IL7 on Ly6D⁺CD135⁺CD127⁺CD19⁻ progenitors, we analyzed a transgenic mouse model in which *Il7* expression is driven by an MHC Class II promoter, resulting in increased *in vivo* levels of IL7 (37). These mice exhibit a lymphoproliferative phenotype with increased numbers of CD19⁺ B cells (38). In contrast to bone marrow CD19⁺ cells, Ly6D⁺ EPLM numbers did not increase in response to elevated IL7 (Figure 5A-C). In addition, the cell cycle profile of Ly6D⁺ EPLM remained unaltered in *Il7*tg mice compared to WT (Figure 5D), arguing against a proliferative action of IL7 on these progenitors. To exclude the possibility that a proliferative signal by FL present in these mice compromised the effect of increased IL7 on the cell cycle status of Ly6D⁺ EPLM, we crossed *Il7*tg with *Flt3l⁻/⁻* mice. Over-expression of IL7 *in vivo* did not result in a significant increase in Ly6D⁺ EPLM or CLP numbers in the absence of FL (Figure 5E and FigS8). In contrast, a 3-fold increase in CD19⁺CD117⁺ numbers was observed (Figure 5F), in agreement with the proliferative effect of IL7 on CD19⁺ B cells. This resulted in a small, but significant, increase in splenic follicular B cells (Fig. S9). Moreover, cell cycle analysis of *il7*tg-*Flt3l⁻/⁻* Ly6D⁺ EPLM showed no significant change in their cycling profile compared to their *Flt3l⁻/⁻* counterparts (Figure 5G). Therefore, we conclude that while IL7 acts as a proliferative factor for CD19⁺ committed B cells, it does not do so for their Ly6D⁺CD135⁺CD127⁺CD19⁻ precursors.

**FL induces proliferation of Ly6D⁺CD135⁺CD127⁺CD19⁻ progenitors**

As evident in Figure 5G, loss of *in vivo* FL signaling affected the proliferative status of Ly6D⁺ EPLM. Comparison of Ly6D⁺ EPLM numbers in mice either lacking or over-expressing FL showed a 14-fold reduction in *Flt3l⁻/⁻* Ly6D⁺ EPLM numbers compared to WT, while *Flt3Ltg* Ly6D⁺ EPLM increased 105-fold (Figure 6A). A similar response to FL levels was observed for Ly6D⁺ CLP (Figure 6A). Cell cycle analysis of Ly6D⁺ EPLM from these mice showed a significant increase in the percentage of Ki67⁻ DAPI⁻ cells and a decrease in the percentage of Ki67⁺ cells when FL signaling was absent, while *Flt3Ltg* Ly6D⁺ EPLM showed the reverse (Figure 6B and Fig. S10). Thus,
our data indicate that FL promotes the proliferation of Ly6D^+CD135^+CD127^-CD19^- progenitors.

To evaluate if FL also regulates the survival of Ly6D^+CD135^+CD127^-CD19^- progenitors, we crossed Flt3l^-/- mice with Bcl2tg mice. Thus, Bcl2tg-Flt3l^-/- mice showed a minor 2-fold increase in Ly6D^+ EPLM numbers compared to their Flt3l^-/- counterparts (1.8-fold for Ly6D^+ CLP) (Figure 6C). Nevertheless, the in vitro B cell potential of Flt3l^-/- Ly6D^+ EPLM progenitors was not improved by Bcl2 overexpression (Figure 6D). Downstream CD19^+ progenitors also demonstrated a partial, but significant, rescue (Figure 6E). Our analysis of Bcl2tg-Flt3l^-/- mice suggests that the reduction in Ly6D^+CD135^+CD127^-CD19^- progenitors observed in Flt3l^-/- mice can only be partially explained by a survival role of FL. In contrast, the clear change in the numbers and cycling profile of these progenitors in response to the absence or overabundance of FL in vivo, as well as the inability of Bcl2 to rescue their in vitro B cell potential, points towards proliferation as being the main effector function of FL at this developmental stage.

**FL does not instruct commitment to the B cell lineage**

The striking rescue in B cell commitment observed in our Flt3Ltg-Il7^-/- mice could be explained by a potential instructive role of FL when present at high levels in vivo. However, increased FL did not result in Ebf1 or Pax5 up-regulation (Figure 3B-D). Moreover, analysis of Flt3l^-/- Ly6D^+ EPLM showed that while absence of FL in vivo leads to a reduction in the numbers of Ly6D^+ EPLM (Figure 1C), it does not significantly reduce the percentage of Ebf1^+ cells within the population (Figure 7A-B), consistent with a permissive rather than instructive role of FL. Finally, the decrease in the Ebf1^+ fraction of Ly6D^+ EPLM upon exposure to high levels of FL was reflected in the increased ability of these progenitors to give rise to T cells in vitro, as manifested by the high frequency of T cell clone generation when Flt3Ltg Ly6D^+ EPLM were plated on OP9DL1 stromal cells in the presence of IL7 (Figure 7C). The above data suggest that FL does not instruct commitment to the B cell lineage through up-regulation of Ebf1 and Pax5 expression.
DISCUSSION

Commitment to the B cell lineage is mediated by the expression of Ebf1 and Pax5 transcription factors and it is initiated in CD135⁺CD127⁺Ly6D⁺ progenitors prior to CD19 expression (30, 31). In il7⁻/⁻ mice, this Ly6D⁺ CLP compartment is significantly reduced (22), a finding confirmed in the present study for both CLP and EPLM, a B220int/+ population partly overlapping with CLP and pre-pro B cells (Figure 1B-C). The proliferative effect of IL7 on committed CD19⁺ B cell progenitors (38) makes the investigation of its role in B cell commitment challenging when using CD19⁺ cells as readout. Hence, we assessed the role of IL7 in B cell commitment by analyzing the Ly6D⁺ CLP/EPLM compartment in different mouse models. Our analysis of Flt3Ltg-il7⁻/⁻ mice showed a complete rescue of Ly6D⁺ CLP/EPLM numbers in vivo and their B cell potential in vitro and in vivo, while Ebf1 and Pax5 were expressed at similar levels to Flt3Ltg mice, thereby indicating that IL7 signaling is not required for their up-regulation at the Ly6D⁺CD19⁻ stage (Figures 2 and 3). These results suggest that IL7 is not acting as an instructive cytokine in B cell commitment by initiating Ebf1 and Pax5 expression at the CD135⁺CD127⁺CD19⁻ stage, as previously hypothesized (18-20), but rather as a permissive one.

Early investigations had shown that Bcl2 over-expression in the absence of IL7 signaling could rescue T cell (39, 40) but not B cell development (16, 17). However, a more recent study demonstrated a Bcl2-mediated rescue of CD19⁺ progenitors in conditional Stat5⁻/⁻ mice, as well as a strong activation of the pro-survival gene Mcl1 expression by Stat5 (21), therefore suggesting a survival role for IL7 in B cell development. Our use of il7⁻/⁻ mice instead of il7ra⁻/⁻, which allows the assessment of progenitor in vitro B cell potential, and our focus on Ly6D⁺CD135⁺CD127⁺CD19⁻ progenitors, has enabled us to confirm the latter findings and extend them to the CD19⁻ stage where B cell commitment events are initiated at the molecular level. Interestingly, il7tg mice analysis showed that IL7 indeed acts as a proliferative factor for committed CD19⁺ cells, but not for their CD19⁻ precursors. Even in the absence of FL, excess IL7 was unable to significantly increase Ly6D⁺ CLP/EPLM numbers, while it did so for CD19⁺ B cell progenitors (Figure 5). Hence, we propose that the main role of IL7 at the CD135⁺CD127⁺CD19⁻ stage is to provide
survival signals to the progenitors until they commit to the B cell lineage upon Pax5 and CD19 expression, after which it additionally induces their proliferation (Figure 7D). This survival role becomes particularly critical when FL levels are limiting, thereby explaining the reduction in Ly6D⁺ CLP/EPLM seen in Il7⁻/⁻ mice. Our study, in agreement with previous data (21), identifies a common, permissive rather than instructive role for IL7 in both B and T cell development (39, 40).

The rescue in B cell commitment without active IL7 signaling occurs when FL is expressed above physiological levels. Even though a minor role for FL as a survival factor for CD135⁺CD127⁺Ly6D⁺CD19⁻ progenitors cannot be excluded, the main effect of FL on these progenitors seems to be the induction of their proliferation, as suggested by their expansion and their increased cycling upon FL over-expression, with the reverse phenotype observed upon loss of FL signaling (Figure 6). Moreover, increased FL leads to expansion of Lin⁻CD117⁺Sca1⁺ cells (LSK) (8), thereby increasing the developmental input into the CD135⁺CD127⁺Ly6D⁺CD19⁻ progenitor stage. None of the mouse models analyzed in the present study gave any evidence for an instructive role of FL in B cell commitment. In contrast, excess FL resulted in a proportional reduction of Ebf- and Pax5-expressing Ly6D⁺CD19⁻ progenitors (Figures 3 and 7). One explanation for this reduction could be the increased percentage of cycling Flt3Ltg Ly6D⁺CD19⁻ progenitors, resulting in a decreased fraction initiating the B cell developmental program. Alternatively, another environmental factor, responsible for initiation of Ebf1/Pax5 expression and B cell commitment, could be the limiting factor in Flt3Ltg mice, thus leading to a smaller fraction of the expanded Ly6D⁺CD19⁻ compartment entering the B cell pathway. Our conclusion is that FL is mainly responsible for generating enough CD135⁺CD127⁺Ly6D⁺CD19⁻ progenitors, both by inducing their proliferation and by increasing their developmental input from the LSK compartment (Figure 7D) (41, 42). As a result, increased levels of FL in Flt3Ltg-Ill7⁻/⁻ mice lead to a dramatic increase in CD135⁺CD127⁺Ly6D⁺CD19⁻ progenitor numbers, therefore surpassing the need for the survival role of IL7 at this stage and resulting in a sufficient fraction of them committing to the B cell lineage.

The generation of B cell progenitors in Flt3Ltg-Ill7⁻/⁻ mice is reminiscent of the apparent IL7 independency of human B lymphopoiesis, where relatively normal numbers of B cells are seen in patients with mutations in components of the IL7
signaling pathway (13, 14). However, all patients with such mutations are neonates and in neonatal II7− mice, B cell development also takes place (43). Therefore, the apparent difference in the IL7 dependency of B cell development between man and mouse could actually reflect the corresponding difference between fetal/neonatal and adult lymphopoiesis. Our data showing that increased FL signaling can rescue B cell commitment in the absence of IL7 could provide a potential explanation for this difference. Fetal/neonatal CD135+CD127−CD19− progenitors might be exposed to higher levels of FL and/or show higher sensitivity to FL signaling than adult CD135+CD127+CD19− progenitors. Indeed, previous studies showed that despite a preferable response of fetal B cell progenitors to TSLP, FL signaling remains an absolute requirement for fetal B lymphopoiesis (44, 45).

The instructive or permissive progenitor regulation of lineage commitment by cytokines is a complex process, in which cytokines can initiate developmental transcription programs in progenitors. However, the reverse is also true, since the particular epigenetic, transcriptional and signaling landscape of a cell can affect its response to a cytokine (9). Indeed, while previous analysis of Flt3Ltg mice indicated an instructive role for FL in promoting differentiation of multi-potent progenitors towards lympho-myeloid and away from erythroid fate (8), our present data show that FL acts in a permissive manner for B cell commitment of CD135+CD127+CD19− progenitors. In addition, whereas IL7 induces proliferation of committed CD19+ B cell progenitors, it does not do so on CD127+CD19− progenitors, suggesting that upon commitment to the B cell lineage, changes in the transcription factor and intracellular signaling landscape influence the effector function of IL7. Therefore, our present data further support the notion of a cell-context dependent cytokine action.

The Ebf1/Pax5 up-regulation and subsequent B cell commitment in Flt3Ltg-II7− mice shown herein raises the issue of the potential extra-cellular regulation of B cell commitment. One possibility could be that another environmental signal from the bone marrow microenvironment - other than IL7, TSLP and FL - initiates Ebf1 expression in CD135+CD127+Ly6D+CD19− progenitors resulting in Pax5/CD19 expression and B cell commitment. Alternatively, as yet uncommitted CD135+CD127+Ly6D+CD19− progenitors could express Ebf1 in a cell-autonomous, stochastic, manner with some obtaining sufficient Ebf1 to initiate the B cell gene
program and eventually commit to the B cell lineage. The intricate transcription factor network sustaining B cell commitment through a series of positive feedback regulatory loops (46) provides conceptual support for the latter hypothesis.
MATERIALS AND METHODS

Mice
For breeding and analysis, age- and sex-matched C57BL/6 Flt3l−/− (27), Flt3Ltg (8), Il7−/− (10), Il7ra−/− (11), Il7tg (38), and (C57BL/6 x C3H) Bcl2tg (35) mice backcrossed with C57BL/6 for at least 5 generations were used at 6–11 weeks of age. All mice were bred and maintained in our animal facility under specific pathogen-free conditions. Animal experiments were carried out within institutional guidelines (authorization number 1888 from cantonal veterinarian office, Basel).

Antibodies, flow cytometry, and sorting
For analysis, cells were flushed from femurs of the two hind legs of mice. The procedure was performed in PBS containing 0.5% BSA and 5mM EDTA. For detection of Ebf1 and cell cycle analysis, cells were fixed and permeabilized after cell-surface staining using the Foxp3 Fix/Perm buffer set (eBioscience), and subsequently stained with PE-conjugated anti-Ebf1 (T26-818) or FITC-conjugated anti-Ki67 (B56) and DAPI, according to the supplier’s protocol. Flow cytometry was done using a BD LSRFortessa (BD Biosciences) and data were analyzed using FlowJo Software (Treestar). For cell sorting, a FACSaria Ilu (BD Biosciences) was used (>98% purity).

In vitro limiting dilution assays
Experiments have been performed as previously described (47). Briefly, OP9 or OP9DL1 stromal cells were plated on flat-bottom 96-well plates one day before the initiation of co-cultures, at a concentration of 3000 cells per well. The following day stromal cells were γ-irradiated (3000 rad) and the sorted progenitor cells were added at different concentrations. Cultures were maintained in IMDM medium supplemented with 5 × 10^{-5} M β-mercaptoethanol, 1 mM glutamine, 0.03% (wt/vol) primatone, 100 U/mL penicillin, 100 μg/mL streptomycin, 5% FBS and 10% IL7-conditioned medium. After 14 days in culture all wells were inspected under an inverted microscope and wells containing colonies of more than 50 cells were scored as positive.
Quantitative real-time PCR
RNA extraction was performed using TRI Reagent® (Life Technologies) followed by cDNA synthesis using GoScript™ Reverse Transcriptase (Promega). Real-time PCR was performed using SYBR Green PCR Master Mix (Applied Biosystems).

Statistical analysis
Statistical analysis was performed with Prism 6.0g software (GraphPad Software, Inc.). Two-tailed unpaired Student t tests were used for statistical comparisons. If not differently indicated, data are presented as mean values ± SD or SEM. n.s. not significant or P > 0.05, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001.

AUTHOR CONTRIBUTIONS
L.v.M. performed experiments, analyzed data and revised the manuscript; L.A-S., F.K. and G.C. performed experiments; D.F. provided Il7-/ and Il7rα-/- mice and revised the manuscript; R.C provided Il7tg mice and revised the manuscript; A.R. designed experiments, analyzed data, revised the manuscript and supervised the project; P.T. designed and performed experiments, analyzed data, wrote the manuscript and supervised the project.
REFERENCES

17. Maraskovsky E, Peschon JJ, McKenna H, Teepe M, & Strasser A (1998) Overexpression of Bcl-2 does not rescue impaired B lymphopoiesis in IL-7
receptor-deficient mice but can enhance survival of mature B cells. *International immunology* 10(9):1367-1375.


Figure 1. IL7 and FL are necessary for the generation of a normal Ly6D⁺CD135⁺CD127⁻CD19⁻ compartment. A. FACS plots showing the gating strategy used for identification of Ly6D⁺ EPLM and their percentage of CD135 and CD127 expression. Lineage staining: SiglecH, CD115, CD11c, NK1.1, Gr-1. B. Representative FACS plots of EPLM (upper row) and CLP (lower row) from the bone marrow of WT, II7⁻/⁻ and Flt3l⁻/⁻ mice. C. Absolute numbers of Ly6D⁺ EPLM (upper graph) and CLP (lower graph) from the bone marrow of WT (n=13), II7⁻/⁻ (n=5) and Flt3l⁻/⁻ (n=10) mice. D. Representative FACS plots of EPLM and CLP from WT and Flt3Ltg mice. E. Absolute numbers of total EPLM and CLP (left graphs) and Ly6D⁺ EPLM and CLP (right graphs) from WT and Flt3Ltg mice.
Figure 2. Increased in vivo FL levels rescue B cell generation in Il7−/− mice. A. Representative FACS plots of EPLM (upper panel) and CLP (lower panel) from WT, Il7−/−, Flt3Ltg and Flt3Ltg-Il7−/− mice. B. Numbers of EPLM (top left), CLP (bottom left), Ly6D+ EPLM (top right) and Ly6D+ CLP (bottom right) from the mouse genotypes indicated on the x-axes. For each mouse genotype mean±SEM is shown. C. Numbers of CD19+CD117+ (top left), CD19+CD117+IgM− (top right) and CD19+IgM+ (bottom) bone marrow cells from the mice indicated on the x-axes. For each mouse genotype mean±SEM is shown. D. Numbers of CD19+CD21highCD23low marginal zone (left) and CD19+CD21+CD23+ follicular (right) B cells in the spleens of WT or mutant mice, as indicated on the x-axes. For each mouse genotype mean±SD is shown.
Figure 3. Increased *in vivo* FL rescues B cell commitment in the absence of IL7 and/or TSLP. **A.** *In vitro* limiting dilution analysis of Ly6D⁺ EPLM B cell potential. Ly6D⁺ EPLM were sorted from WT, II7⁻⁻, Flt3Ltg and Flt3Ltg-II7⁻⁻ mice and plated at the indicated concentrations on OP9 stromal cells together with IL7. One representative out of four independent experiments is shown. **B.** Real-time quantitative PCR analysis showing expression of Ebf1, Pax5 and Foxo1 mRNAs in Ly6D⁺ EPLM sorted from the indicated mouse genotypes. Bars show fold expression
relative to WT (set as 1). Error bars represent the SEM from 3-6 independent experiments. C. Representative FACS plots showing expression of Ebf1 protein within the Ly6D+ EPLM of the indicated WT or mutant mice. D. Percentages of Ebf1-expressing Ly6D+ EPLM from WT (n=7), Il7−/− (n=3), Flt3Ltg (n=11) and Flt3Ltg-II7−/− (n=6) mice. Bars show mean±SEM. E. Ly6D+ EPLM (left) and CD19−CD117+ (right) numbers from WT (n=5), Il7−/− (n=5), Flt3Ltg (n=3), Flt3Ltg-II7−/− (n=5) mice, as well as from Il7−/− (n=5) and Il7rα−/− (n=6) mice injected intra-peritoneally with 10 daily doses of 10 μg FL each (indicated as +FL) or PBS (+PBS, n=4). Shown is the mean±SEM.
Figure 4. Bcl2 over-expression partially rescues B cell commitment in Il7−/− mice. A. Representative FACS plots of EPLM (upper panel) and CLP (lower panel) from WT, Il7−/−, Bcl2tg and Bcl2tg-il7−/− mice. B. Numbers of EPLM (top left), CLP (bottom left), Ly6D+ EPLM (top right) and Ly6D+ CLP (bottom right) from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown. C. In vitro limiting dilution analysis of Ly6D+ EPLM B cell potential. Ly6D+ EPLM were sorted from WT, Il7−/−, Bcl2tg and Bcl2tg-il7−/− mice and plated at the indicated concentrations on OP9 stromal cells together with IL7. One representative out of three independent experiments is shown. D. Numbers of CD19+CD117+ (top), CD19+CD117+IgM− (middle) and CD19+IgM+ (bottom) bone marrow cells from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown.
Figure 5. IL7 does not induce proliferation of Ly6D⁺CD135⁺CD127⁺CD19⁻ progenitors. A. CD19⁺CD117⁺ numbers in bone marrow of WT (n=10), II7⁻⁻ (n=5) and II7tg (n=8) mice. B. EPLM numbers in bone marrow of WT (n=14), II7⁻⁻ (n=7) and II7tg (n=5) mice. C. Ly6D⁺ EPLM numbers in bone marrow of WT (n=14), II7⁻⁻ (n=7) and II7tg (n=5) mice. D. Cell cycle analysis of Ly6D⁺ EPLM from WT (n=5) and II7tg (n=2) mice. Graph shows percentages of Ki67⁻⁻DAPI⁻, Ki67⁺DAPI⁻ and Ki67⁺DAPI⁺ Ly6D⁺ EPLM. Bars in A, B, C and D show mean±SEM. E. Numbers of EPLM (top left), CLP (bottom left), Ly6D⁺ EPLM (top right) and Ly6D⁺ CLP (bottom right) from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown. F. Numbers of CD19⁺CD117⁺ bone marrow cells from WT and mutant mice, as indicated on the x-axis. For each mouse genotype mean±SEM is shown. G. Cell cycle analysis of Ly6D⁺ EPLM from WT (n=5), Flt3⁻⁻ (n=3), II7tg (n=2) and II7tg-Flt3⁻⁻ (n=3) mice. Graph shows percentages of Ki67⁻⁻DAPI⁻, Ki67⁺DAPI⁻ and Ki67⁺DAPI⁺ Ly6D⁺ EPLM. Bars show mean±SEM.
Figure 6. Fl promotes proliferation but not survival of Ly6D+CD135+CD127–CD19– progenitors. A. Numbers of EPLM (top left), CLP (bottom left), Ly6D+ EPLM (top right) and Ly6D+ CLP (bottom right) from WT (n=14), Flt3l−/− (n=10) and Flt3Ltg (n=9) mice. Bars show mean±SEM. B. Cell cycle analysis of Ly6D+ EPLM from WT (n=5), Flt3l−/− (n=3) and Flt3Ltg (n=9) mice. Graph shows percentages of Ki67−DAPI−, Ki67−DAPI− and Ki67+DAPI+ Ly6D+ EPLM. Bars show mean±SEM. C. Numbers of EPLM (top left), CLP (bottom left), Ly6D+ EPLM (top right) and Ly6D+ CLP (bottom right) from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown. D. In vitro limiting dilution analysis of Ly6D+ EPLM B cell potential. Ly6D+ EPLM were sorted from WT, Flt3l−/− and Bcl2tg-Flt3l−/− mice and plated at the indicated concentrations on OP9 stromal cells together with IL7. E. Numbers of CD19+CD117+ (left), CD19+CD117+IgM− (middle) and CD19+IgM+ (right) bone marrow cells from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown.
Figure 7. FL does not instruct *Ebf1* expression and B cell commitment. A. Representative FACS plots showing expression of *Ebf1* protein within the Ly6D⁺ EPLM of WT, *Flt3l*⁻/⁻ and *Flt3Ltg* mice. B. Percentages of *Ebf1*-expressing Ly6D⁺ EPLM from WT (n=7), *Flt3l*⁻/⁻ (n=5) and *Flt3Ltg* (n=12) mice. Bars show mean±SEM. C. *In vitro* limiting dilution analysis of Ly6D⁺ EPLM T cell potential. Ly6D⁺ EPLM were sorted from WT and *Flt3Ltg* mice and plated at the indicated concentrations on OP9DL1 stromal cells together with IL7. One representative of four independent experiments is shown. D. Schematic model for the permissive role of IL7 and FL acting on hematopoietic progenitors and CD19⁺ committed B-cell precursors. HSC: Hematopoietic Stem Cell; LMPP: Lymphoid-primed Multi-Potent Progenitor; CLP: Common Lymphoid Progenitor; EPLM: Early Progenitor with Lymphoid and Myeloid potential.
Supporting Materials and Methods

**Antibodies**

The following antibodies were used for flow cytometry (from BD Pharmingen, eBioscience, BioLegend, or produced in house): anti-B220 (RA3-6B2), anti-CD117 (2B8), anti-CD19 (1D3), anti-NK1.1 (PK136), anti-SiglecH (551), anti-CD11c (HL3), anti-CD115 (AFS98), anti-Ly6D (49-H4), anti-CD127 (SB/199), anti-CD135 (A2F10), anti-Sca1 (D7), anti-IgM (M41), anti-CD21 (7G6), anti-CD23 (B3B4), anti-CD4 (GK1.5), anti-CD8 (53.6.7), anti-TCRβ (H57).

**Quantitative real-time PCR**

The primers used were: *Ebf1*: Ebf1-F: 5’-CAGGAAACCCACGTGACAT-3’; Ebf1-R: 5’-CCACGTTGACTGTGGTAGACA-3’, *Pax5*: Pax5-F: 5’ACGCTGACAGGGATGGTG-3’; Pax5-R: 5’-GGGGAACCTCCAAGAATCAT-3’, *Foxo1*: Foxo1-F: 5’-AGTGGATGGTGGAAGACGT-3’, Foxo1-R: 5’-GAAGGGACAGATTGTGGCG-3’, *Actin*: Actin-F: 5’-CTGTCGAGTCCGTCCACC-3’, Actin-R: 5’-CGCAGCGATATCGTATCCA-3’.
**Fig. S1.** A. CLP FACS staining in WT mice. FACS plots showing the gating strategy used for the identification of Ly6D⁺ CLP. Lineage staining: SiglecH, CD115, CD11c, NK1.1, Gr-1. B. *In vitro* limiting dilution analysis of Ly6D⁺ and Ly6D⁻ EPLM B cell potential. Cells were sorted as shown in Figure 1A and plated at the indicated concentrations on OP9 stromal cells together with IL7. A representative of 3 independent experiments is shown. C, D. Numbers of EPLM (C) and CLP (D) progenitors in WT (n=13), Il7⁻/⁻ (n=5) and Flt3l⁻/⁻ (n=10) mice. EPLM were stained as shown in Figure 1A and CLP as shown in A. Student’s t test. ***P ≤ 0.001. Bars show mean±SEM.
Fig. S2. Rescue of CD19<sup>+</sup> bone marrow B cell progenitors in Flt3ltg-I<sup>−/−</sup> mice. Figure shows representative FACS plots for the identification of CD19<sup>+</sup>CD117<sup>+</sup>, CD19<sup>+</sup>CD117<sup>−</sup>IgM<sup>−</sup> and CD19<sup>+</sup>IgM<sup>+</sup> bone marrow cells.
Fig. S3. Thymic T cell development in Flt3tg-II7^{-/-} mice. A. Representative FACS plots showing CD4/CD8 thymocyte staining from 6-8 week old WT, II7^{-/-}, Flt3tg and Flt3tg-II7^{-/-} mice (n=4 for each group). B. Total numbers of CD4^{+} (left panel) and CD8^{+} (right panel) single-positive thymocytes from the mouse genotypes indicated on the x-axes.
Fig. S4. **A.** Representative FACS plots illustrating T cells in the spleens of WT (first row), II7<sup>−/−</sup> (second row), Flt3Ltg (third row), and Flt3Ltg-II7<sup>−/−</sup> (fourth row) mice. After gating on living lymphocytes TCRβ<sup>+</sup> cells are further sub-grouped in CD4 and CD8 positive T cells. **B.** Representative FACS plots illustrating B cells in the spleens of WT (first row), II7<sup>−/−</sup> (second row), Flt3Ltg (third row), and Flt3Ltg-II7<sup>−/−</sup> (fourth row) mice. After gating on living lymphocytes CD19<sup>+</sup> cells are further sub-grouped in CD21<sup>high</sup>CD23<sup>low</sup> marginal zone B cells and CD21<sup>+</sup>CD23<sup>+</sup> follicular B cells. **C** and **D.** Numbers of splenic CD4<sup>+</sup> (C) and CD8<sup>+</sup> (D) T cells, stained as shown in A, from WT and mutant mice as indicated on the x-axes. ***P ≤ 0.001, ****P ≤ 0.0001. Student’s t test; n = 9-15. Data shown above are mean ± SD.
**Fig. S5.** B cell potential of Ly6D\(^+\) EPLM cells from \(\text{iI}7^{-/-}\) mice injected with FL. \(\text{iI}7^{-/-}\) mice were injected with FL (10 daily doses of 10\(\mu\)g for each mouse) and Ly6D\(^+\) EPLM were sorted from their bone marrows one day after the last injection. **A.** In vitro limiting dilution analysis of the B cell potential of FL-injected \(\text{iI}7^{-/-}\) Ly6D\(^+\) EPLM. Cells were plated at the indicated concentrations on OP9 stromal cells plus IL7. Flt3Ltg Ly6D\(^+\) EPLM were used as positive controls. **B, C.** In vivo B cell potential of FL-injected \(\text{iI}7^{-/-}\) Ly6D\(^+\) EPLM. Five thousand Ly6D\(^+\) EPLM from FL-injected \(\text{iI}7^{-/-}\) or Flt3Ltg mice were intravenously injected into sub-lethally irradiated \(\text{Rag}2^{-/-}\) mice. Four weeks after cell transfer spleens were analyzed for expression of CD19 and IgM. **B.** Respresentative FACS plots of recipient spleens. **C.** Numbers of CD19\(^+\)IgM\(^+\) B cells harvested from the analyzed spleens (n=4 mice for each group).
Fig. S6. Quiescent state of Bcl2-rescued cells in vivo. A. Cell cycle analysis of Ly6D$^+$ EPLM from WT (n=5), Bcl2tg, (n=2), Il7$^{-/-}$ (n=2) and Bcl2tg-Il7$^{-/-}$ (n=4) mice. Graph shows percentages of Ki67$^-$DAPI$^-$, Ki67$^+$DAPI$^-$ and Ki67$^+$DAPI$^+$ Ly6D$^+$ EPLM. *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001. Student’s t test. Bars show mean±SEM. B. Representative Ki67/DAPI FACS plots of the Ly6D$^+$ EPLM cell cycle analysis collectively presented in A.
Fig. S7. Bcl2-mediated rescue of splenic T and B cells in the absence of IL7. A. Representative FACS plots illustrating T cells in the spleens of WT (first row), \( \text{Il7}^{-/-} \) (second row), \( \text{Bcl2tg} \) (third row), and \( \text{Bcl2tg-Il7}^{-/-} \) (fourth row) mice. After gating on living lymphocytes TCR\( \beta \)^+ cells are further sub-grouped in CD4 and CD8 positive T
cells. **B.** Representative FACS plots illustrating B cells in the spleens of WT (first row), *Il7*^-/-* (second row), Bcl2tg (third row), and Bcl2tg-Il7^-/-* (fourth row) mice. After gating on living lymphocytes CD19^+^ cells are further sub-grouped in CD21^{high}CD23^{low} marginal zone B cells and CD21^+CD23^+ follicular B cells. **C.** Numbers of splenic CD4^+^ (top) and CD8^+^ (bottom) T cells, stained as shown in **A**, from WT and mutant mice as indicated on the x-axes. **D.** Numbers of splenic marginal zone (top) and follicular (bottom) B cells, stained as shown in **B**, from WT and mutant mice as indicated on the x-axes. **P \leq 0.01, ****P \leq 0.0001. Student’s t test; n = 4-9. Data shown above are mean ± SD.
Fig. S8. Representative FACS plots of EPLM (upper panel) and CLP (lower panel) from WT, Flt3−/−, Il7tg and Il7tg-Flt3−/− mice.
Fig. S9. Effect of IL7 over-expression on WT and Flt3l^{-/-} splenic T and B cells. A. Representative FACS plots illustrating T cells in the spleens of WT (first row), Flt3l^{-/-} (second row), Il7tg (third row), and Il7tg-Flt3l^{-/-} (fourth row) mice. After gating on living lymphocytes TCRβ^{+} cells are further sub-grouped in CD4 and CD8 positive T cells. B. Representative FACS plots illustrating B cells in the spleens of WT (first row),
Flt3" (second row), Il7tg (third row), and Il7tg-Flt3" (fourth row) mice. After gating on living lymphocytes CD19+ cells are further sub-grouped in CD21^{\text{high}}CD23^{\text{low}} marginal zone B cells and CD21^{+}CD23^{+} follicular B cells. C. Numbers of splenic CD4^{+} (top) and CD8^{+} (bottom) T cells, stained as shown in A, from WT and mutant mice as indicated on the x-axes. D. Numbers of splenic marginal zone (top) and follicular (bottom) B cells, stained as shown in B, from WT and mutant mice as indicated on the x-axes. ns not significant or P > 0.05, ****P ≤ 0.0001. Student’s t test; n = 3-15. Data shown above are mean ± SD.
Fig. S10. Effect of in vivo FL levels on Ly6D$^+$ EPLM cell cycle. Representative Ki67/DAPI FACS plots of the Ly6D$^+$ EPLM cell cycle analysis collectively presented in Figure 6B.
9.2 Paper 2:

Reconstitution of a functional B-cell compartment in immunodeficient mice
with pro-B cells propagated with or without stromal cells

Lilly von Muenchow\textsuperscript{1}, Panagiotis Tsapogas\textsuperscript{3}, Llucia Albertí-Servera\textsuperscript{1}, Giuseppina Capoferr\textsuperscript{1}, Marianne Doelz\textsuperscript{1}, Hannie Rolink\textsuperscript{1}, Nabil Bosco\textsuperscript{1,2}, Rhodri Ceredig\textsuperscript{3} and Antonius G. Rolink\textsuperscript{1}

\textsuperscript{1} Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland

\textsuperscript{2} Present address: Nestlé Research Center, Lausanne, Switzerland

\textsuperscript{3} Department of Biosciences, University of Galway, Galway, Ireland

Correspondence: Prof. Antonius G. Rolink; e-mail: antonius.rolink@unibas.ch

Acknowledgements

A.G.R. is holder of the chair in immunology endowed by L. Hoffmann – La Roche Ltd, Basel. This study was supported by the Swiss National Science Foundation and by the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013 under Research Executive Agency grant agreement number 315902. R.C. was supported by Science Foundation Ireland under grant numbers SFI09/SRC/B1794 and SFI07/SK/B1233b.

Conflict of interest

The authors declare no commercial or financial conflict of interest.
ABSTRACT

Fetal liver (FL) and bone marrow (BM) derived pro-B cells were propagated long-term in stromal cell free cultures supplemented with interleukin-7 (IL-7), stem cell factor and FLT3 ligand. Within a week, most cells expressed surface CD19, CD79A, λ5 and VpreB antigens and had rearranged immunoglobulin D-J heavy chain genes. Both FL and BM pro-B cells reconstituted the B-cell compartments of immuno-incompetent Rag2-deficient mice with FL pro-B generating follicular, marginal zone (MZB) and B1a B cells, but BM pro-B cells mainly MZB. Reconstituted mice generated significant IgM and IgG antibodies to a type II T-independent antigen with FL pro-B cell reconstituted mice generating surprisingly high IgG1 titers. Finally, we show for the first time that mice reconstituted with mixtures of stromal cell free \textit{in vitro} propagated pro-B and pro-T cells mounted a T-cell dependent antibody response. This novel stromal cell free culture system facilitates our understanding of B-cell development and might be applied clinically.
INTRODUCTION

In mammals, B cells develop from hematopoietic stem cells (HSCs), antenatally from the fetal liver (FL) and postnatally from the bone marrow (BM). Various stages of B-cell development are distinguishable by combinations of cell surface and intracellular markers, cell cycle profile and rearrangement status of IgH and IgL genes [1-4]. B-cell commitment is determined by the transcription factor Pax5 [5-7] one of whose target genes is CD19 [8]. The earliest B-cell committed precursor, or pro-B cell, proliferates rapidly, is CD19^+CD117^+ [1, 9] and has its IgM D-J genes rearranged [4] and are absent in Pax5-deficient [10], Il7 or Il7R gene deleted mice [11, 12]. The receptor tyrosine kinases CD117 and CD135 and their corresponding ligands, SCF and FLT3L, are important for early B-cell development [13-17]. Blocking SCF binding with an anti-CD117 antibody inhibits pro-B cell proliferation in IL-7-containing stromal cell cultures [18] and CD135 or FLT3L-deficient mice have a dramatically reduced BM pro-B-cell compartment [19]. Thus Pax5 and cytokines are important for early B-cell development.

We previously showed that FL-derived pro-B cells could be grown long-term on stromal cells and IL-7 [20]. Here we present for the first time that this is possible for BM-derived pro-B cells. We have described an early (E) progenitor (P) with lymphoid (L) and myeloid (M) developmental potential, called EPLM, in the BM [21]. EPLM are B220^+, CD117^+ but CD19^− NK1.1^-. When cultured either on OP9 stromal cells plus IL-7 or without stromal cells but with IL-7, SCF and FLT3L, EPLM from the FL or BM can be propagated long-term and differentiate into CD19^+ pro-B cells. Moreover, upon in vivo transplantation into immunodeficient Rag2 gene deficient hosts, these pro-B cells reconstitute a functional B-cell compartment. Finally, when these mice are reconstituted with a mixture of in vitro stromal cell free propagated pro-B and pro-T cells [22], a small, but functional adaptive immune system is generated.
RESULTS

Growth of BM pro-B cells long-term

Using stromal cell-based cultures, only FL-derived pro-B cells could be grown long-term [20]. We recently identified a B220−CD117−CD19−NK1.1− BM progenitor having lymphoid and myeloid developmental potential, called EPLM [21]; now we tested their capacity to generate long-term growing pro-B cells. Thus, B220−CD117−CD19−NK1.1− (CD19−) and B220−CD117−CD19+NK1.1− (CD19+) cells from FL or BM were sorted (Fig.1A) and plated on OP9 stromal cells plus IL-7. After 6 days, cells were harvested every 3 - 4 days and re-plated on fresh stromal cells and IL-7. FACS staining 6-7 days after initiation of cultures showed in all cases that the cultured cells were >90% positive for CD19 expression and this expression was further increased to almost 100% and retained throughout the culture period (data not shown and Figure 1D). CD19− (EPLM) cells grew continuously with doubling times of about 30hrs, whereas CD19+ cells proliferated for the first 7 days and then died (Fig 1B). RNAseq analysis on CD19+ cells derived from CD19− EPLM grown for 12 days on OP9 plus IL-7 and freshly-isolated CD19+CD117+ cells similarly grown for 5 days revealed that only 83 genes, none of which were B-cell related, were ≥2 fold differentially expressed (Table SI). All 44 B cell related genes were similarly expressed (Table SII). Gene ontology analysis of the 83 genes did not identify genes that explained the growth difference observed. An additional RNAseq analysis of ex vivo isolated CD19+CD117+ cells and EPLM derived CD19+ cells maintained on OP9 stromal cells in the presence of IL-7 for two weeks revealed 3449 differentially expressed genes (Table SIII). However, amongst them we could not identify particular candidates that might be responsible for the dissimilar growth capacity observed.

FL-derived EPLM grew better than BM-derived EPLM on OP9 with IL-7 (Fig.1C) with >95% of both becoming CD19+ (Fig.1D) with ~25% BM and >60% FL-derived cells expressing CD117 (Fig.1D). RT-PCR analysis revealed that both expressed CD79a, CD79b, Igll1 (λ5) and Vpreb1 (Fig.1E) and both had undergone D\textsubscript{H}-J\textsubscript{H} rearrangements (Fig.1F). Thus, BM and FL EPLM cultured on OP9 plus IL-7 give rise to pro-B cells proliferating for more than three months.
**Long-term propagation of BM and FL pro-B cells in stromal cell free cultures**

FL pro-B cells propagated on stromal cells plus IL-7 reconstituted the B-cell compartment of immuno-deficient mice [20, 29, 30]. However, it is unlikely that pro-B cells derived from co-culture settings will ever be approved for therapeutic purposes. Therefore, we developed a stromal cell free culture system consisting of soluble IL-7, SCF and FLT3L. As shown in figure 2A and B, EPLM from FL and BM showed very robust growth under these stromal cell free conditions. Indeed, the growth rate of BM-derived EPLM was identical under stromal cell and stromal cell free conditions (Fig.2A). However, growth of FL EPLM was slightly slower without stromal-cell support (Fig.2B). Stromal cell free cultured EPLM became CD19+ with ~35% BM and >75% FL-derived cells expressing CD117 (Fig.2C and D). Moreover, cells cultured without stroma expressed CD79a and b, Igll1 (λ5) and Vpreb1 (Fig.2E) and had their IgH chains D_{H-JH} rearranged (Fig.2F). Thus, both BM and FL EPLM cultured with IL-7, SCF and FLT3L alone also give rise to long-term proliferating pro-B cells. Moreover, as shown for FL-derived pro-B cells [20], BM-derived pro-B cells expressing the anti-apoptotic Bcl2 transgene cultured with or without stromal-cell support efficiently differentiated into IgM+ B cells upon IL-7 removal (Figure S1).

**In vivo B-cell reconstitution by BM or FL EPLM-derived pro-B cells.**

To test whether *in vitro*-generated FL or BM-derived pro-B cells could reconstitute mice, 10^7 EPLM-derived pro-B from FL or BM of B6 CD45.1 mice cultured with or without stroma were transferred into sub-lethally irradiated CD45.2 B6 Rag2-deficient mice. After 5-10 weeks, FACS analysis of spleen or peritoneal cavity (PerC) cells of FL (Fig 3A) or BM (Fig 4A) reconstituted mice showed that expression of donor CD45.1+ was restricted to CD19+ B cells (first column) and in both cases, all CD19+ cells were IgM+ (second column). In figure 3, WT cytograms are shown in the first row. Based on CD5 expression, FL-derived pro-B cells propagated with or without stroma gave rise to a large fraction of B1a B cells especially in the PerC (Fig.3A third column cytograms). As expected, BM-derived pro-B cells generated few CD5+ B cells (Fig.4A third column cytograms). Using combined expression of CD21 and CD23 to define CD21+/CD23- marginal zone B (MZB) and
CD21+/CD23+ follicular B cells (FB) spleen CD19+ B cells derived from FL pro-B cells propagated with or without stroma were similar (Fig3A fourth column cytograms). Thus 40 – 60% were FB and 30 – 40% were MZB. The CD19+/CD21+/CD23- cells (lower left quadrant) most likely represent B1 B cells. Immunohistochemical analysis (Fig 3B) of reconstituted mice showed the typical B cell follicular structure seen in WT mice comprising an outer ring of IgMhigh (green) IgDlow (blue) MZB cells surrounding metallophilic macrophages (red) with IgMpositive IgDhigh FB inside. Mice reconstituted with BM pro-B cells cultured by the two methods also showed no obvious differences in CD21 and CD23 expression (Fig4A fourth column cytograms). However, unlike FL-derived pro-B cells, >70% of BM pro-B cell-derived B cells were MZB and only 15 – 20% FB (Fig3B fourth column cytograms).

Spleens of FL-derived pro-B cell reconstituted mice contained around 5 x 10^6 B cells with no difference between cells cultured with or without stroma (Fig.4B). In contrast, spleens of mice reconstituted with BM-derived pro-B cells contained only 0.5 x 10^6 B cells irrespective of whether they had been propagated with or without stroma. Thus FL-derived pro-B cells seem to be much more efficient at reconstitution than BM-derived cells.

To test this more stringently, competitive reconstitution experiments were performed. Sub-lethally irradiated CD45.2 B6 Rag2-deficient mice were reconstituted with a 1:1 mixture of 5 x 10^6 CD45.2 FL-derived and CD45.1 BM-derived pro-B cells. After 8 weeks, FACS analysis of one representative mouse (Fig 4C) showed that 44% of splenocytes were CD19+ of which 97.5% (42.9/44) were FL and 2.5% (1.1/4.4) BM-derived. Similar results were obtained in more than three independent experiments. Thus FL-derived pro-B cells are superior to BM-derived ones also in competitive transplantation settings.

In order to identify genes that might be responsible for these observed differences between FL- and BM-derived pro-B cells we performed RNA-sequencing analysis of the two populations. This analysis identified 218 genes differentially expressed more than 2-fold and with high significance (Table SIV). We identified the gene Lin28B, as the one most highly expressed in FL-derived pro-B cells compared to BM-derived pro-B (34-fold). It was recently shown that enforced Lin28B gene expression in adult HSC converted them functionally into FL HSC [31] and that BM-
derived pro-B cells gained FL-derived pro-B cell properties [32, 33]. To test whether Lin28B expression also influenced reconstitution efficiency, we introduced a MigR1 retrovirus encoding Lin28B into BM-derived pro-B cells propagated on OP9 stromal cells for two weeks. Then $5 \times 10^6$ Lin28B transduced and $5 \times 10^6$ non-transduced BM-derived pro-B cells were co-transferred into sub-lethally irradiated Rag2-deficient recipients. Only 2% CD19$^+$ cells were found in the spleens of these mice of which half expressed Lin28B (GFP) (Figure S2). Thus Lin28B expression did not improve the in vivo B-cell generating capacity of BM-derived pro-B cells. However, practically all CD19$^+$ cells were IgM$^+$ and > 80% Lin28B$^+$ cells were CD5$^+$, thereby resembling B1 B cells. Concerning CD5 expression, Lin28B over-expression also converted BM pro-B cells to ones phenotypically resembling FL cells.

**Pro-B cell derived B cells mount a T-cell independent immune response**

To test whether the B-cell compartments of Rag2-deficient reconstituted mice were functional, 8 weeks after cell transfer they were immunized with the T-cell independent antigen NIP-Ficoll. Serum anti-NIP titers were determined one week before and two weeks after immunization. All reconstituted mice mounted a good IgM anti-NIP response with IgM titers comparable to those in immunized wild type B6 mice (Fig.5A). Mice reconstituted with BM-derived pro-B cells showed a rather low, but significant, IgG anti-NIP response (Fig.5B). However, the IgG anti-NIP response of mice reconstituted with FL pro-B cells was as high or even higher than wild type B6 mice (Fig.5B). Thus the pro-B cell reconstituted B-cell compartments were functional.

The surprisingly high IgG anti-NIP titers in FL pro-B reconstituted mice prompted us to determine their isotypes. Thus, the IgG$_{2A}$ and IgG$_{2B}$ anti-NIP titers were low (Fig 5C) and comparable to those in immunized wild type B6 mice whereas IgG$_{1}$ and IgG$_{3}$ titers were considerably higher. Although high IgG$_{3}$ titers are observed in other T-cell independent responses, switching to IgG$_{1}$ was thought to be a highly T-cell dependent phenomenon requiring IL-4 [34-36]. In addition to T cells, mast cells [37, 38], basophils [39], eosinophils [40] and ILC2s [41] may produce IL-4. ILC generation is largely dependent on IL-7 and Rag2/common gamma chain (Rag2cγ) double-deficient mice are practically devoid of ILCs [42, 43]. Therefore, Rag2 and
Rag2cγ double-deficient mice were reconstituted with $10^7$ FL-derived pro-B cells and immune responses analyzed as above. Both Rag2 and Rag2cγ double-deficient mice showed a very significant IgM and IgG anti-NIP response (Fig. 6A and B), similar and/or even higher than wild type B6 mice. Moreover, both types of reconstituted mice showed a very strong IgG$_1$ and IgG$_3$ anti-NIP response (Fig. 6C and D). Thus the IgG class switching observed in immunized reconstituted mice does not seem to be regulated by ILCs.

**Reconstitution of the adaptive immune system by in vitro-propagated pro-B cells and pro-T cells**

Recently, we described a stromal cell free culture system for the long-term propagation of pro-T cells that could be used to reconstitute the T-cell compartment of T-cell deficient mice [22]. However, reconstituted mice contained few regulatory T cells and developed a wasting disease preventable by the co-transfer of mature Treg cells or co-transfer of pro-T cells transduced with a retrovirus encoding Foxp3-IRESCRE-GFP. Simultaneous reconstitution with a mixture of in vitro-propagated pro-B cells and a non-transduced and Foxp3 transduced pro-T cells in a 4:1 ratio resulted in T but no B-cell reconstitution (not shown). Therefore, sub-lethally irradiated B6 Rag2-deficient mice were first reconstituted with $10^7$ FL-derived pro-B cells and 4 weeks later with $10^7$ pro-T cells of which $2.5 \times 10^6$ were Foxp3$^+$. Six weeks after pro-T cell-transfer, a significant B and T-cell reconstitution was seen in peripheral blood cells (data not shown). Because FL pro-B cells partially reconstituted the B-cell compartment of B-cell deficient, T-cell proficient, µMt mice and mounted a T-cell dependent immune response (Figure S3) we tested whether the established B and T cells were functional and could cooperate. Reconstituted mice were therefore immunized with the T-cell dependent antigen NIP-OVA. Rag2-deficient mice reconstituted with pro-B cells alone were used as controls and serum anti-NIP titers determined one week before and two weeks after immunization. No IgM or IgG anti-NIP response was observed in control pro-B cell-reconstituted Rag2-deficient mice (Fig. 7A). However, mice reconstituted with both pro-B and pro-T cells mounted a strong IgM and relatively weak, but significant, IgG anti-NIP response (Fig. 7A). Thus the reconstituted adaptive immune system was functional.
At 10 – 14 weeks after pro-T cell-transfer, the extent of reconstitution was assessed by FACS analysis. In one such mouse, spleen lymphocytes comprised 10% B cells, 10% CD8 T cells and 20% CD4 T cells (Fig.7B) with ~10% CD4 T cells expressing GFP, indicating they were Tregs. Spleen cell numbers of 7 individual mice 14 weeks after pro-T cell-transfer (Fig.7C) show they contained around 3.5 x 10^6 CD19^+ IgM^+ B cells, 1 x 10^6 CD8 T cells and 3 x 10^6 CD4 T cells of which 0.5 x 10^6 were GFP positive; derived from Foxp3 transduced pro-T cells. Thus transfer of in vitro-propagated pro-B and pro-T cells into Rag2-deficient mice resulted into reconstitution of a small, but functionally active, adaptive immune compartment.
DISCUSSION

To our knowledge, this is the first report describing the long-term propagation of BM-derived pro-B cells and their use in reconstituting a functional immune system in immune-deficient recipients. Pro-B cells from Pax5−/− mice were previously shown to have multi-lineage developmental potential [44] and we identified an equivalent B220+CD117+CD19−NK1.1− EPLM cell with lymphoid and myeloid developmental potential (called EPLM) in the BM of WT mice [21]. Herein we show that BM-derived EPLM cultured on OP9 stromal cells plus IL-7 differentiated into CD19+, CD79a+, CD79b+, Igll1+ and Vpreb1+ pro-B cells with rearranged D_{H}-J_{H} genes. Unlike freshly isolated CD19+ CD117+ pro-B cells, those derived from EPLM could be cultured long-term in vitro. Gene expression profiling did not reveal an obvious explanation for this difference in growth capacity but could possibly be regulated via RNA modification or at the translational level [45-47].

To date, long-term growth of pro-B cells required co-culture on stromal cells, yet the specific role of stromal cells in this culture system was unknown. Clearly IL-7, IL-7R (CD127), SCF and its receptor (CD117) as well as FLT3 and its receptor (CD135) all play a role in B-cell development and/or pro-B cell growth [11, 12, 15, 16, 18, 19]. Indeed, we now show that EPLMs efficiently grow and differentiate into pro-B cells when cultured in the combined presence of IL-7, SCF and FLT3L without contact with stromal cells. This indirectly suggests that SCF and FLT3 can substitute for stromal cells.

Pro-B cells maintained on stromal cells plus IL-7 could reconstitute the B-cell compartment of immunodeficient mice [20] and herein we extend these findings. Both FL and BM EPLM-derived pro-B cells propagated either on stromal cells or stromal cell free generated a significant B-cell compartment upon transfer into Rag2-deficient mice. However, FL-derived pro-B cells were about 40 fold more efficient at B-cell reconstitution than their BM-derived partners.

Although pro-B cells expressing the Lin28B can acquire characteristics of FL-derived pro-B cells [32] and in vitro-propagated Lin28B-expressing BM-derived pro-B cells could generate CD5 positive B cells in vivo, there was no improvement in the in vivo-reconstitution capacity of BM pro-B cells.
An enlarged MZB compartment is frequently observed in B lymphopenic mice [48-52] suggesting that newly-formed B cells first fill the MZB compartment. This phenomenon could be due either to a specialized environment within the splenic marginal zone or that only B cells expressing certain BCR home to and expand in this anatomical location [53].

When immunized with the T-independent antigen NIP-Ficoll, all reconstituted mice mounted an anti-NIP response indicating their B cell compartments were functional. Mice reconstituted with BM-derived pro-B cells had a slightly lower IgM anti-NIP titer than WT mice possibly due to poor B cell reconstitution and showed a significant IgG anti-NIP response although still at least 10 fold lower than WT mice. In marked contrast, Rag2-deficient mice reconstituted with FL pro-B cells mounted an IgM and IgG anti-NIP response that was identical, or even higher than, WT mice suggesting that fetal progenitor-derived B cells are the main anti-NIP responders in a primary immunization. In B6 mice, the primary anti-NIP antibodies use predominantly λ1 light chains [54-56]. The anti-NIP response of pro-B cell-reconstituted mice was likewise dominated by λ1-containing antibodies (data not shown) indicating that the anti-NIP B-cell repertoires in pro-B cell reconstituted mice are identical to those of WT mice.

Surprisingly, FL pro-B cell-reconstituted mice had a high IgG1 anti-NIP response an antibody class normally associated with IL-4. Since reconstituted mice were devoid of T cells, the question arose which cell type was responsible for IL-4 production [34-36]. Pro-B cell reconstituted Rag2cy double-deficient mice, deficient in ILCs, also mounted high IgG1 response suggesting that ILCs were not the source of IL-4 [42, 43]. Mast cells, basophils and/or eosinophils, could be the source of IL-4 [37-40]. FL pro-B cells were also able to reconstitute a functional B-cell compartment in T-cell containing, B-cell deficient μMt mice (Figure S3). This result might seem surprising given that the T-cell compartments in these mice may not be tolerant to mature B cells.

Pro-T cells propagated in vitro under stromal free conditions can reconstitute the T-cell compartments of T-cell deficient mice [22]. Here we show that the combined transfer of stromal cell free-propagated pro-B and pro-T cells into Rag2-deficient mice results in the generation of a functional adaptive immune repertoire
capable of mounting a T-dependent antibody response. For efficient reconstitution, pro-T cell transfer had to be performed 2-3 weeks after the pro-B cell reconstitution suggesting that when transferred together, T cells develop that are not tolerant to B cells and therefore eliminate them.

Overall the findings described herein show that progenitor lymphocytes can be readily propagated under stromal free conditions in vitro and that these cells can be used to reconstitute mice with mature functional lymphocytes. Based on these results the establishment of stromal cell free culture systems for human lymphocyte progenitors might be of great interest since unlike stromal cell-propagated lymphocyte progenitors, stromal cell free propagated cells could be potentially used for therapeutic purposes in patients with B and/or T-cell deficiencies.
MATERIALS AND METHODS

Mice
Female C57BL/6 CD45.1 and CD45.2, C57BL/6 Rag2-deficient [23], and C57BL/6 Bcl-2 transgenic mice [24] with 5–8 weeks of age were used. The appearance of vaginal plugs was counted as day 0 of gestation and embryos were taken at day E17.5. All mice were bred and maintained in our animal facility under specific pathogen-free conditions. Animal experiments were carried out within institutional guidelines (authorization numbers 1886 and 1888 from Kantonales Veterinäramt, Basel).

Cell lines, cell culture, and supplements
The OP9 stromal cell line [25] was cultured as a monolayer in IMDM supplemented with 2% FBS, 5 × 10⁻⁵ M β-mercaptoethanol, 1 mM glutamine, 0.03% w/v Primatone (Quest, Naarden, The Netherlands), and 100 U/mL penicillin. For pro-B cell culture, CD117⁺B220⁺CD19⁻NK1.1⁻ cells were sorted from the FL or the femoral BM of adult mice and cultured at 10⁴/ml in supplemented IMDM either on a semi-confluent layer of 30 Gy γ-irradiated OP9 stromal cells in the presence of 100U/ml IL-7, or without stromal cell support but in the presence of 100U/ml IL-7, 50ng/ml FLT3L, and 100ng/ml SCF. IL-7 was derived from culture supernatant of J558L cells transfected with murine IL-7 cDNA. Polyhistidin-tagged SCF was purified from transfected Rosetta pLacl bacteria using Ni-NTA-agarose beads (Qiagen, Venlo, NL). A vector expressing a human FLT3L-Fc fusion protein was expressed in Chinese hamster ovary cells. The supernatant was passed over a protein A-Sepharose (GE Healthcare, Chalfont St. Giles, GB) column in order to purify the protein. Pro T cells were cultured as previously described [22].

Antibodies, flow cytometry, and sorting
FITC-, PE-, allophycocyanin-, or biotin-labeled mAbs specific for CD117, B220, CD19, NK1.1, IgM, Igκ, CD5, CD45.1, CD21, CD23, CD4, CD8α, and TCRβ were either purchased from BD Biosciences (Franklin Lakes, NJ, USA) or eBiosciences (San Diego, CA, USA), or purified from hybridoma culture supernatants according to standard procedures. Staining of the cells was performed as described before [26]. Flow
cytometry was done using a FACS Calibur (BD Biosciences) and data were analyzed using the CellQuest Pro (BD Biosciences) or FlowJo Software (Treestar). For cell sorting, a FACSAria Ilu (BD Biosciences) was used (>98% purity).

**Transfer of cultured progenitor cells**

Recipient mice were γ-irradiated using a Cobalt source (Gammacell 40, Atomic Energy of Canada, Ltd) 4h prior to reconstitution. The indicated number of pro-B or pro T cells was then injected into the tail vein.

**Immunohistochemical analysis**

To analyze pro-B-cell derived B-cell localization in the spleen, the 5µm snap frozen and acetone-fixed sections were incubated with FITC-labelled anti-IgM (clone M41, self-made), APC-labelled anti-IgD (clone 1.19 self-made), and biotinylated anti-MOMA-1 (Vector, Burlingame, CA), which was revealed with PE-conjugated streptavidin (SouthernBiotech, Birmingham, AL). Confocal microscopy images were taken with a LSM 510 Meta (Zeiss, Oberkochen, D) and analyzed using the ImageJ software and the Fiji image processing package.

**NIP-specific antibody responses**

Reconstituted mice were immunized subcutaneously with 100μg NIP-Ficoll or NIP-OVA in a 1:1 CFA emulsion. Serum IgM and IgG against NIP was analyzed at day 14 using ELISA as described in [27].

**PCR analyses**

The PCR conditions for amplifying D_{i}H_{i} rearrangements were described elsewhere [4, 6, 28]. The primers used were D_{i} 5´-TTCAAGCACAATGCTGGCT-3´ and J_{i} 5´-GTCTAGATTCTCACAAGAGTCCGATAGACCTGG-3´. Oligonucleotide primers used for CD79A verification from pro-B cDNA were 5´-TGTTTTGGGTCCCAGTGAGCA-3´ and 5´-CACGGAGGGAAGTACCACA-3´, for CD79B 5´-TCTTTCAGGTAGCAGCCGGA-3´ and 5´-TATGTGTCGGCTGTACAT-3´, for IGLL1 5´-AGTAGGACAGACTCTGGCA-3´ and 5´-GGCTGACCTAGGATGTGACG-3´, for VPREB1 5´-CTCCGGGTCGAAAGATACGAC-3´ and 5´-GCTCATAGCAACCCCGAGAA-3´, and for beta-actin
5’-GAAGTCTAGAGCAACATAGCACAGCTTCTC-3’ and
5’-GTGGGAATTCGTCAGAAGGACTCCTATGTG-3’.

**Statistical analysis**

Statistical analysis was performed with Prism 6.0g software (GraphPad Software, Inc.). Two-tailed unpaired Student t tests were used for statistical comparisons. If not differently indicated, data are presented as mean values ± SEM from three independent experiments. ns not significant or P > 0.05, * P ≤ 0.05, ** P ≤ 0.01, *** P ≤ 0.001, ****P ≤ 0.0001.
REFERENCES


15 Mackarehtschian, K., Hardin, J. D., Moore, K. A., Boast, S., Goff, S. P. and Lemischka, I. R., Targeted disruption of the flk2/flt3 gene leads to


Doherty, T. A., Khorram, N., Lund, S., Mehta, A. K., Croft, M. and Broide, D. H., Lung type 2 innate lymphoid cells express cysteinyl leukotriene receptor 1,


Figure 1. Establishment of EPLM-derived long-term proliferating pro-B cell lines from BM and FL. FL and BM-derived CD19⁺B220⁺CD117⁺NK1.1⁻ cells were sorted and maintained on OP9 stromal cells plus IL-7 (A) Gating strategy for sorting of CD19⁺ and CD19⁻ B220⁺CD117⁺ BM cells. (B) In vitro growth capacity of sorted CD19⁺ and CD19⁻ BM cells on OP9 stromal cells plus IL-7. (C) Comparison of the in vitro growth of FL and BM-derived pro-B cells for about 40 days. (D) Representative FACS plots showing B220, CD19 and CD117 expression in FL and BM-derived pro-B cells cultured for 12 days. (E) RT-PCR analysis of Cd79a, Cd79b, Igll1, Vpreb1, and Actb expression in FL and BM-derived pro-B cells cultured for 12 days. (F) Genomic D_H-J_H rearrangement analysis of FL and BM-derived pro-B cells cultures for 12 days.
Figure 2. Establishment of EPLM-derived long-term proliferating pro-B cell lines from BM and FL under stromal cell free conditions. BM or FL-derived CD19<sup>-</sup> B220<sup>-</sup>CD117<sup>+</sup>NK1.1<sup>-</sup> cells were sorted and maintained either on OP9 stromal cells plus IL-7 or in the presence of IL-7, SCF and FLT3L without stromal cells. (A) Comparison of the growth capacity of BM-derived pro-B cells cultured in the absence (black squares) or presence (white circles) of stromal cells for about 40 days. (B) Comparison of the growth capacity of FL-derived pro-B cells cultured in the absence (black squares) or presence (white circles) of stromal cells for about 40 days. (C) Representative FACS plots showing B220, CD19 and CD117 expression on BM-
derived pro-B cells cultured for 14 days with or without stromal cells, as indicated. (D) Representative FACS plots showing B220, CD19 and CD117 expression on FL-derived pro-B cells cultured for 14 days with or without stromal cells, as indicated. (E) RT-PCR analysis of *Cd79a*, *Cd78b*, *Igll1*, *Vpreb1* and *Actb* expression in FL and BM-derived pro-B cells cultured for 14 days in the presence of IL-7, SCF and FLT3L without stromal cells. (F) Genomic D_{H}-J_{H} rearrangement analysis of FL and BM-derived pro-B cells cultured for 14 days in the presence of IL-7, SCF and FLT3L without stromal cells.
Figure 3. In vivo reconstitution potential of FL-derived pro-B cells propagated in vitro. Sub-lethally irradiated CD45.2 B6 Rag2-deficient mice were injected intravenously with $10^7$ CD45.1 FL-derived pro-B cells propagated in the presence or absence of stromal cells. (A) Representative FACS analysis of B cell populations in the spleen and peritoneal cavity (PerC) of WT controls and recipient mice 8 weeks after cell transfer of the indicated pro-B cells. (B) Representative picture of a staining for IgM (green), IgD (blue), and metallophilic macrophages (MOMA-1 in red) on a spleen section of mice reconstituted with FL-derived pro-B cells propagated on OP9 stromal cells (20x magnification). In total 28 individual mice were injected: 14 mice with FL-derived pro-B cells propagated with OP9 stromal cells and 14 mice with FL-derived pro-B cells propagated without stromal cells.
Figure 4. In vivo reconstitution potential of BM-derived pro-B cells propagated in vitro and comparison to the corresponding potential of FL-derived pro-B cells. Sub-lethally irradiated CD45.2 B6 Rag2-deficient mice (n=21) were injected intravenously with $10^7$ CD45.1 BM-derived pro-B cells propagated in the presence (14 mice injected) or absence (7 mice injected) of stromal cells. (A) Representative FACS analysis of B-cell populations in the spleen and peritoneal cavity (PerC) of WT controls and recipient mice 8 weeks after cell transfer of the indicated pro-B cells. (B) Absolute numbers of CD19$^+$ IgM$^+$ cells in the spleens of mice injected with BM-derived pro-B cells and FL-derived pro-B cells (Figure 3). Unpaired student’s t-test. Data shown above are mean ± SEM and each circle or square is an individual mouse.
ns: not significant or P > 0.05, **P ≤ 0.01 and ****P ≤ 0.0001. (C) CD19, IgM, and CD45.1 expression from a spleen of a CD45.2 B6 Rag2-deficient mouse 8 weeks after transfer of 5 x 10^6 CD45.2 FL-derived pro-B cells and 5 x 10^6 CD45.1 BM-derived pro-B cells. Both pro-B cells were propagated stromal cell free. A representative FACS analysis out of 3 independently performed experiments is shown.
Figure 5. T-cell independent responses of mice reconstituted with pro-B cells propagated in the absence or presence of stromal cells. B6 Rag2-deficient mice were reconstituted with BM or FL derived pro-B cells propagated with or without stromal cells and immunized with NIP-Ficoll 8 weeks after cell transfer. Normal B6 (WT) mice were used as controls and statistical significance compared to WT is indicated below the after immunization data points (black squares). Titers were defined as the serum dilutions that gave 2 times background OD values in the ELISA test. Titers of the indicated antibodies were determined in sera taken 1 week before immunization (A and B open symbols) and 2 weeks after immunization (A and B closed symbols). (A) IgM anti-NIP titers. (B) IgG anti-NIP titers. (C) Anti-NIP IgG subclass titers in the serum of mice reconstituted with FL-derived pro-B cells and WT mice. ns: not significant or P > 0.05, * or * P ≤ 0.05, ** or ++ P ≤ 0.01, *** or +++ P ≤ 0.001, **** P ≤ 0.0001. * represents significance compared to pre-bleed. + represents significance compared to titer of WT mice. Each symbol represents an individual mouse. Unpaired student’s t test; n = 5-8. Data shown above are mean ± SEM and two independent experiments were performed.
Figure 6. Anti-NIP responses of WT mice, Rag2-deficient and Rag2cγ double-deficient mice reconstituted with FL pro-B cells propagated on stromal cells and immunized with NIP-Ficoll 8 weeks after cell transfer. Titers were defined as the serum dilutions that gave 2 times background OD values in the ELISA test. Titers were determined in sera taken 1 week before immunization (A and B open symbols) and 2 weeks after immunization (A and B closed symbols) (A) IgM anti-NIP titers. (B) IgG anti-NIP titers. (C) IgG1 anti-NIP titers. Sera from B6 mice immunized with NIP-Ficoll or NIP-OVA were used as positive controls. (D) IgG3 anti-NIP titers. Sera from B6 mice immunized with NIP-Ficoll or NIP-OVA were used as positive controls. Collective data from two independent experiments are shown and each symbol represents an individual mouse. ns: not significant, * P ≤ 0.05, ** P ≤ 0.01, *** P ≤ 0.001, **** P ≤ 0.0001. Unpaired student’s t test; n = 5-8. Data shown above are mean ± SEM.
Figure 7. In vivo lymphoid reconstitution and immune responses in mice injected with in vitro propagated pro-B and pro-T cells. Sublethally irradiated Rag2-deficient mice were reconstituted with FL-derived pro-B cells and 4 weeks later with a 3:1 mixture of in vitro propagated pro-T cells [22] and Foxp3-transduced in vitro propagated pro-T cells. Mice were immunized with NIP-OVA 11 weeks after pro-B cell transfer and splenocytes analyzed by flow cytometry 3-7 weeks later. (A) IgM and IgG anti-NIP responses of reconstituted and immunized B6 Rag2-deficient mice. Sera were collected 1 week before and 2 weeks after immunization. (B) Representative FACS analysis of splenic B- and T-cell populations 10-14 weeks after pro-T cell transfer. (C) Absolute numbers of T and B cells found in the spleens of reconstituted mice. ns: not significant or $P > 0.05$, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. Unpaired student’s t test; n = 7-9. Numbers above data points indicate mean ± SEM. Two independent experiments were performed and each symbol represents an individual mouse.
Figure S1
BM EPLM-derived pro-B cells propagated either on stromal cells or grown stromal-cell free differentiated efficiently into IgM positive B cells upon IL-7 withdrawal. Shown is μH and Igk expression by BM-derived CD19'B220'CD117'NK1.1' cells from Bcl2 transgenic mice cultured on OP9 stromal cells plus IL-7 or cultured in the presence of IL-7, SCF and FLT3L (upper dot blots) and by the same cells cultured for 3 weeks as described above and then for 6 days in plain medium. Whereas in the presence of IL-7 very few cells expressed a μH (IgM) and Igk light chain on their surface, withdrawal of IL-7 resulted in 26% of pro-B cells cultured on stromal cells becoming IgM+, Igk+. The remaining 7.1% IgM+, Igk- (upper left quadrant) cells being Igκ- (data not shown). These values were 17.3% and 7.8% respectively for pro-B cells cultured in stromal cell-free conditions (lower right dot plot). This has been observed in more than 3 independent experiments.
Figure S2
CD19 and GFP expression and IgM, GFP and CD5 expression on CD19-gated cells recovered from B6 Rag2-deficient mice reconstituted with 5 x 10^8 BM-derived pro-B cells and 5 x 10^6 BM-derived Lin28B-IRES-GFP transduced pro-B cells. Analysis was performed 8 weeks after cell transfer. This has been observed in more than three independent experiments.
Figure S3
Reconstitution of the B-cell compartment in μMT mice with FL-derived pro-B cells. (A) CD19 and IgM expression by μMT spleen cells 8 weeks after cell transfer. (B) IgM anti-NIP titer of μMT mice immunized with NIP-OVA at 8-10 weeks after pro-B cell transfer. Anti-NIP titers were determined in the sera taken 2 weeks after immunization. Sera taken 1 week before immunization were used as negative controls. Sera from WT B6 mice immunized with NIP-OVA were used as positive controls. (C) IgG anti-NIP titer of μMT mice immunized with NIP-OVA at 8-10 weeks after pro-B cell transfer. Anti-NIP titers were determined in the sera taken 2 weeks after immunization. Sera taken 1 week before immunization were used as negative controls. Sera from WT B6 mice immunized with NIP-OVA were used as positive controls. *P > 0.05, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. Student’s t test; n = 5. Data shown above are mean ± SEM.
9.3 Review:

Versatility of stem and progenitor cells and the instructive actions of cytokines shape haematopoiesis

Geoffrey Brown¹, Ciaran J Mooney¹, Llucia Alberti-Servera², Lilly von Münchow², Kai-Michael Toellner¹, Rhodri Ceredig³ and Antonius G Rolink²

¹ School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
² Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland
³ Regenerative Medicine Institute, National Centre for Biomedical Engineering Science, Department of Physiology, National University of Ireland, Galway, Ireland
Versatility of stem and progenitor cells and the instructive actions of cytokines on hematopoiesis

Geoffrey Brown1, Ciaran James Mooney1, Lucia Alberti-Servera2, Lilly von Muenchow3, Kai-Michael Toellner1, Rhodri Cerecig2, and Antonius Rolink3

1School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham, UK
2Developmental and Molecular Immunology Laboratory, Department of Biomedicine, University of Basel, Basel, Switzerland, and 3Regenerative Medicine Institute, National Centre for Biomedical Engineering Science, Department of Physiology, National University of Ireland, Galway, Ireland

Abstract
For many years, developing hematopoietic cells have been strictly compartmentalized into a rare population of multi-potent self-renewing hematopoietic stem cells (HSCs), multi-potent hematopoietic progenitor cells (MPPs) that are undergoing commitment to particular lineage fates, and recognizable precursor cells that mature towards functional blood and immune cells. A single route to each end-cell type is prescribed in the “classical” model for the architecture of hematopoiesis. Recent findings have led to the viewpoint that HSCs and MPPs are more versatile than previously thought. Underlying this are multiple routes to a particular fate and cells having clandestine fate options even when they have progressed some way along a pathway. The primary role of cytokines during hematopoiesis has long been seen to be regulation of the survival and proliferation of developing hematopoietic cells. Some cytokines now clearly have instructive actions on cell fate decisions. All this leads to a new way of viewing hematopoiesis whereby versatile HSC and MPP are directed towards lineage outcomes via cytokine regulated cell-fate decisions. This means greater flexibility to the shaping of hematopoiesis.


Introduction
The hematopoietic stem cell (HSC) gives rise to a wide range of blood and immune cell types. The generation of large numbers of each cell type is a complex and tightly regulated process that is ultimately governed by commitment of rare, and generally quiescent, HSCs to pathways of cell differentiation. These cells, which reside within the bone marrow in adult mammals, are the apex of the hematopoietic hierarchy. HSCs, which can self-renew, give rise to multi-potent
progenitor cells (MPPs) that undergo decision-making, expansion, and differentiation, via recognizable lineage precursors, to give rise to the final compartment of functional cells. Since the early 1980s, the end-cell types have been viewed as two families: lymphoid that includes B and T lymphocytes and natural killer (NK) cells, and myeloid consisting of the rest of the blood and immune cells. Commensurate with this has been the identification of progenitors of each family; namely the common lymphoid progenitor (CLP) and common myeloid progenitor (CMP). This lymphoid/myeloid dichotomy is the basis of a long-standing model for the architecture of hematopoiesis which also encompasses single routes of differentiation towards individual fates. Over the past 15 years, the above strict architecture has become less black and white. HSCs and hematopoietic progenitor cells (HPCs) are now viewed as more versatile and developing HSCs/HPCs as less rigorously compartmentalized.

Cell heterogeneity and versatility during hematopoiesis

Heterogeneity of hematopoietic stem cells

HSCs are better described in mouse than in human, and mouse HSCs can be purified to homogeneity to a greater extent. There are differences between mouse and human HSCs, and between fetal and adult HSCs. Human HSCs are isolated for transplantation on the basis of expression of the cell surface molecule CD34. In contrast, a single mouse CD34<sup>lo</sup>-HSC reconstitutes hematopoiesis long-term in a lethally irradiated mouse. Matsuoka et al. observed that HSCs of bone marrow, liver, and spleen from fetal and neonatal mice express CD34, whereas HSCs are enriched in the CD34<sup>+</sup> fraction and MPPs are CD34<sup>-</sup> in mice older than 10 weeks. Albeit, the principles of hematopoiesis derived from human and mouse studies, including of fetal and adult cells, are substantially similar.

The description of HSCs, and isolation by fluorescence activated cell sorting, relies on the presence and absence of a range of cell surface molecules. Studies of HSCs in the mouse have the advantage that HSCs can be rigorously defined during purification as cells that repopulate long-term (LT-HSC) and short-term (ST-HSC) the entire hematopoietic system. A cardinal aspect of HSCs is that they do not express markers that are associated with the various hematopoietic cell lineages (Lin<sup>-</sup>), including, for example, CD3 (T lymphocytes), B220 (B lymphocytes), CD11b (monocytes/macrophages), Ly-6G (neutrophils), and TER-119 (erythroid cells). HSCs and MPPs express the two molecules, c-Ki67, a mast/stem cell growth factor receptor and tyrosine kinase, and Sca-1, a phosphatidylinositol-anchor membrane protein. LT-HSC and ST-HSC reside in the Lin<sup>-</sup>, Sca-1<sup>+</sup>, c-Ki67<sup>+</sup> population of bone marrow cells, termed LSK.

Originally the Weissman group sub-divided the LSK compartment on the basis of expression of Thy-1.1 and the Flt3 (fms-like) tyrosine kinase, a type 3 receptor kinase. The loss of a low level of expression of Thy-1.1 and gain of expression of Flt3 was observed to correlate with a loss of the capacity of HSCs to self-renew (Figure 1). LT-HSCs were isolated from adult bone marrow as LSK Flt3<sup>-</sup>, and transplantation of LSK Flt3<sup>-</sup> cells resulted in short-term multi-lineage reconstitution. However, fetal liver HSCs are contained within the Flt3<sup>+</sup> and Flt3<sup>-</sup> LSK cells. MPPs are Thy1<sup>-</sup> Flt3<sup>+</sup>. In 2005, Yang et al. combined LSK markers with CD34 and Flt3. The definitions provided by the Jacobsen group, which are commonly used, are LT-HSC as LSK CD34<sup>+</sup> Flt3<sup>-</sup>, ST-HSC as LSK CD34<sup>+</sup> Flt3<sup>+</sup>, and MPPs as LSK CD34<sup>-</sup> Flt3<sup>+</sup> (Figure 1).

Recently, the signaling lymphocyte activation molecule (SLAM) family of membrane receptors (CD150, CD48, CD229, and CD244) has been used to describe subpopulations of HSCs and MPPs (Figure 1). Kiel et al. purified HSCs and MPPs as CD150<sup>CD48<sup>CD244</sup> and CD150<sup>-</sup>CD48<sup>-</sup>CD244<sup>-</sup>, respectively. Lineage-restricted progenitors can be identified by acquired expression of CD48 and are mostly CD150<sup>CD48<sup>CD244</sup> and CD150<sup>-</sup>CD48<sup>-</sup>CD244<sup>-</sup>. Adding CD229 to the panel of SLAM, CD34 and LSK markers led Oguro et al. to propose HSC-1 and HSC-2 populations, which have different lineage biases, and three populations of MPPs (Figure 1). CD229<sup>-</sup> HSC-1 cells are cells that rarely divide and are myeloid-biased as revealed by transplantation studies. CD229<sup>+</sup> HSC-2 cells divide more frequently and are lymphoid-biased<sup>110</sup>. Contrary to the findings of the Morrison group, the level of expression of CD150 has also been used to distinguish myeloid- and lymphoid-biased HSCs. Ema et al. have observed that myeloid-biased HSCs are enriched in the CD150<sup>CD48<sup>CD34<sup>+</sup> LSK population and lymphoid-biased HSCs are enriched in CD150<sup>-</sup>CD48<sup>-</sup>CD34<sup>+</sup> LSK population. These workers have proposed that these two populations of cells overlap with LT-HSCs and ST-HSCs.
As mentioned above and for many years, the two compartmentalizing properties of HSCs are their capacities to reconstitute the entire hematopoietic system and to self-renew. Oguro, in summarizing the sub-populations of hematopoietic stem and progenitor cells distinguished using SLAM markers, provides a scheme whereby the self-renewal potential of HSC-1 is long-term, of HSC-2 is long to intermediate, of MPP-1 is intermediate to transient, and of MPP2 and MPP3 is transient. Oguro also described HPCs (HPC-1 and 2) which do not self-renew. In essence, the boundary between HSCs and MPPs as to the property of self-renewal is blurred, and a distinction between HSCs and MPPs is perhaps somewhat redundant. Both these cell types might be best viewed, and classified as a continuum of HPCs that reduce their capacity to self-renew as they mature.

Strict multi-potency does not always go hand in hand with self-renewal as to the identification of HSCs with lineage biases. As mentioned above, the CD34-LSK population of bone marrow cells, which engraft mice including secondary hosts, has been divided into myeloid-biased and lymphoid-biased HSCs. In addition to expressing a higher level of CD150, myeloid-biased HSCs exclude Hoechst 33342 more effectively than lymphoid-biased HSC, and these two HSC sub-types are differentially regulated by transforming growth factor-β1 (TGF-β1)12. There is additional heterogeneity within the CD150+CD34- LSK population of cells, as revealed by engraftment of single cells in irradiated mice. In primary hosts, some single cells readily gave rise to myeloid cells and cells that were able to engraft a secondary host. Some single cells produced few myeloid cells in primary hosts and cells which when transferred to secondary hosts gave rise progressively to multiple lineages13. Furthermore, CD41 and CD86 expression on HSCs has been reported to distinguish myeloid-biased and lymphoid-biased cells, respectively14,15. As mice age, there are quantitative differences in lineage biases in the HSC population. The data support a model whereby myeloid-biased HSCs have clonally expanded, while lymphoid-biased HSCs exhaust themselves due to their more extensive proliferative nature16.

Recently Jacobsen’s group have described a platelet-biased LT-HSC that expresses von-Willebrand factor (vWF)17. Transplantation of single vWF+ HSCs into irradiated hosts resulted in reconstitution biased towards platelets and myeloid cells. vWF+ HSCs gave rise to a lymphoid-biased reconstitution, vWF+ HSCs require thrombopoietin (Tpo) for their maintenance as these cells were significantly reduced in number in Tpo−/− mice18. vWF+ HSCs gave rise to vWF− HSCs, and vWF+ HSCs were not able to give rise to vWF+ HSCs. This led the Jacobsen group to propose that platelet-biased HSCs are the apex of the hematopoietic hierarchy.

Heterogeneity of hematopoietic progenitor cells

In the “classical” lymphoid/myeloid dichotomy model of hematopoiesis, the sets of potentials observed for different types of progenitors align themselves to developmental progression along each arm of the dichotomy. However, this is not the case for all the progenitors that have been described to date. At odds with an irrevocable commitment of HSCs to either a lymphoid or myeloid pathway is the early description of a progenitor in mouse fetal liver with just the potentials for B lymphoid and macrophage differentiation19. This cell was later shown to be present in adult bone marrow20. A further finding that contradicts the notion that HSCs make an immediate and irrevocable decision to commit to either the lymphoid or myeloid pathways of differentiation is the identification of cells with lymphoid potentials and an incomplete set of myeloid potentials. These cells are early progenitors with lymphoid and myeloid potential (EPLM), that can give rise to T and B lymphocytes, NK cells, dendritic cells (DCs), and macrophages21, and lymphoid–primed multi-potent progenitors (LMP), that have little potential for megakaryocyte or erythrocytoid development while retaining other potentials22.

Progenitor cells that contradict a lymphoid/myeloid dichotomy led us to propose the pair-wise model of hematopoiesis23-25. This model does not assume lineage branching patterns or prescribe a single preferred route to a particular end-cell fate. Instead mature cell fates are shown to be near-neighbors within a continuum of lineage fates (Figure 2). As HSCs mature towards a specific dominant cell fate, fates that are distantly related to this fate are lost first and more closely related fates remain possible as latent fates. The model envisages versatility of HSCs and MPPs, as to allowing an end-cell type to be reached by more than one route (see also below). Mapping of transcription factor usage and the responsiveness of progenitors to growth factors support the proposed close relationships between cell lineages22-24.

Figure 2. A pair-wise model of hematopoiesis. A fate choice continuum with an invariant series of pair-wise developmental relationships between hematopoietic cells.
Figure 3. Mapping of hematopoietic progenitor cells to the pair-wise model. The set of fates available to various progenitors underlie near-neighbor placement of lineages in the pair-wise model. Progenitors with commensurate and contiguous lineage potentials are CLP, common lymphoid progenitor; CMP, common myeloid progenitor; DCPro-B, dendritic cell progenitor B; HPC, hematopoietic progenitor; Eo/B-CFU, eosinophil and basophil progenitor; EPLM, early progenitor with lymphoid and myeloid potential; GMP, granulocyte and macrophage progenitor; LMPP, lymphoid-primed multi-potent progenitor; MEP, megakaryocyte and erythrocyte progenitor; Mon/BDC, monocyte, B lymphocyte and dendritic cell progenitor; NK/T, natural killer cell and T lymphocyte progenitor; HSC, hematopoietic stem cell; Ly, lymphoid bias; Meg, megakaryocyte bias; My, myeloid bias.

A number of progenitor cells that have different sets of potentials have been described. The various combinations of lineage potentials that exist within progenitors are also reflected in a number of different cell lines. Figure 3 shows that the different combinations of differentiation capabilities described for normal progenitors can be mapped to the pair-wise model. Of particular importance to placing cell lineages close to one another in the continuum are progenitor cells that have just two differentiation capacities. In other words, the continuum infers that only certain bi-potentialities are permissible. For example, a cell with the potential for megakaryocyte and T cell differentiation should not exist, and has not been described to date. Bi-potent progenitor cells can be placed within the model with the exception of a bi-potent B lymphocyte/macrophage cell. Whether this cell can give rise to DCs has not been studied.

Already there is a considerable variety of stem cells, with differing biases, and progenitor cells with differing sets of potentials. The latter in turn give rise to end-cell types that can be divided into numerous sub-types, for example, as is the case for T helper cells and DCs. The full extent of the heterogeneity of progenitor cell populations, and their mature progeny, is very much contingent on the extent to which new and emerging cell surface markers can be used to define new sub-populations. It is highly likely that progenitor cells with multiple lineage options and that we presently view as a homogeneous population of cells will be divided into cells with lineage biases in various directions.

Versatility of hematopoietic progenitor cells

The pair-wise model allows developmental pathways to be flexible. One aspect of this is HSCs and their progeny using more than one route to a particular end-cell type. This principle was demonstrated by deriving DCs in vivo from cells purified as CLPs and CMPs. When the transcription profiles of the two DC populations were compared they were found to be the same. Ishikawa et al. concluded that the developmental program of human DCs operates independently of the pathways for myeloid and lymphoid cells.

Alongside new findings there has been a plethora of new models for the architecture of hematopoesis. Some of the models depict multiple routes towards granulocytes and monocytes. For example, a model provided by Jacobsen considers the possibilities of these myeloid cells arising from ST-HSC/MPP via (i) CMP and GMP (granulocyte and monocyte progenitor); (ii) LMPP giving rise to GMP, and (iii) LMPP giving rise to a granulocyte/monocyte/T lymphocyte progenitor which in turn gives rise to GMP. In a model proposed by Katsura, the two routes towards myeloid cells are HSCs veering towards: (i) a cell with the potentials for myeloid, erythroid, and megakaryocyte development; and (ii) a cell with the potentials for myeloid and lymphoid development. Ye and Graf compared the production of mature cell types to flows along branches of a tree with a major branch giving rise to platelets, erythroid cells, granulocytes, and monocytes in equal measure and a separate branch giving rise largely to granulocytes and monocytes. Of course, it is difficult to exclude the possibility that all the above routes to myeloid cells occur to some degree.

Precise tracking of a progenitor, as defined by a set of markers, giving rise to the next progenitor, also verified by markers, and so on to an end-cell type is an impossible task. However, a novel approach has been recently described whereby cells can be uniquely marked in situ, by transposon tagging, allowing longitudinal analysis of clones in mice and revealing that long-lived progenitors are a main driver of steady-state hematopoesis in adulthood. This approach will also impact on our view of cell lineages and provide information on the origin of hematopoietic malignancies. Aside from this experimental model, examination of the sets of options available to various progenitors, as defined by appropriate markers, and which progenitors are or are not able via loss of a fate option(s) to give rise to one another, allows configuration of routes to end-cell types. For certain cell types, it is clear that there can be multiple routes. As described by Jacobsen et al., HSCs can give rise to neutrophils and monocytes through LMPP and CMP intermediates. Figure 4 shows different possible routes downstream of LMPP and CMP towards neutrophils. These include...
potential pathways through an eosinophil/granulocyte/monocyte progenitor\(^{32}\) and/or GM-CFU (granulocyte/macrophage colony-forming unit). The routes are in keeping with the pairwise model as to their close proximity. Similarly, there are multiple potential pathways towards monocytes via CMP, an eosinophil/granulocyte/monocyte progenitor and GM-CFU, and via LMPP, EPLM, and a monocyte/dendritic cell progenitor\(^{33}\). In fetal liver, a myeloid/B-biased progenitor is a further intermediary to myeloid cells\(^{34}\). A caveat to all this is the extent to which routes are used in vivo. However, the point of interest is the inherent flexibility of pathways available to progenitors. It is noteworthy that the ability of precursor cells to follow alternative developmental pathways to give rise to the same cell phenotype was described for cell lineages in the embryo of the leech as early as 1987\(^{35}\).

One aspect underpinning versatility is that progenitor cells retain fate options even after they have progressed some way along a pathway. Thymus-settling progenitors (TSP) have the potential to give rise to myeloid cells, dendritic cells, NK cells, and B lymphocytes in addition to T lymphocytes. TSP give rise to double-negative (DN) 1 early thymocyte progenitors (ETP), which give rise to DN2 cells. DN2 cells have lost the potential for differentiation towards B lymphocytes, but when cultured in the right environment generate myeloid, dendritic cells, and NK cells\(^{36-41}\). As to driving pro-T cells to other fates, culture of DN2 cells in IL3 (interleukin) and stem cell factor (SCF) in the absence of Notch signaling revealed mast cell potential, indicating a close unexpected relationship between the pro-T cell and mast cell fates\(^{42}\). These clandestine potentials are lost as DN2 cells progress to the DN3 stage of development.

Some cytokines have instructive actions on cell-fate decisions

Early glimpses to the instructive action of cytokines

A long-standing debate is whether the commitment of HSCs to fate options occurs in a cell-autonomous and stochastic way, or is driven (in an ordered way) by instructive signals from the local environment\(^{43-46}\). Cytokines are the pivotal external factors that impart environmental signals to control hematopoietic cell development. They have multiple actions that can be viewed as either instructive, by directing HSC/MPPs towards a specific lineage, or permissive, by selectively allowing cells committed to a particular lineage to survive and/or proliferate\(^{47,48}\). For many years, a permissive role of cytokines has been favored. A very recent and complete turnabout in our understanding of the control of hematopoiesis is the notion that cytokines instruct decision-making\(^{49-50}\).

Information to support an instructive role for cytokines has been available for quite some time. In 1982, Metcalf and Burgess concluded that granulocyte/macrophage colony-stimulating (GM-CSF) factor and macrophage colony-stimulating factor (M-CSF) can “irreversibly commit the progeny of GM-CFC respectively to granulocyte and macrophage production”\(^{51}\). When paired daughter cells of GM-CFU were split and one cultured in GM-CSF and the other in M-CSF, some of the cells underwent irreversible commitment to the granulocyte and macrophage pathways, respectively. This occurred during completion of the first cell division and within 24 h. Later in 1991, Metcalf again concluded that colony-stimulating factors have the ability to influence lineage commitment\(^{52}\). Metcalf examined the relative frequencies of lineage committed progenitors when blast cell colonies were established from normal bone marrow cells in combinations with granulocyte colony-stimulating factor (G-CSF), GM-CSF, and multi-CSF. The relative frequency of granulocyte progenitors was increased when cultures were established in the combination with GM-CSF or multi-CSF with SCF.

In 2000, Kondo et al. provided even more evidence to support the notion that cytokines can convert the fate of lymphoid-committed progenitors\(^{53}\). The IL-2 and GM-CSF receptors were exogenously expressed in CLPs, which normally give rise exclusively to T lymphocytes, B lymphocytes, and NK cells. This resulted in cell-fate conversion to the myeloid lineage. The use of mutants of the beta-chain of the IL-2 receptor revealed that signals for the granulocyte and monocyte differentiation pathways are provoked by different cytoplasmic domains of the IL-2 receptor. Kondo et al. also showed that primitive HSCs express low-to-moderate levels
of the receptors for GM-CSF and M-CSF. Hence, there is a possibility that HSCs are receptive to the instructive actions of these growth factors. Kondo et al. concluded from all the above that down-regulation of cytokine receptors that drive myeloid cell development is a critical step in commitment of cells to lymphoid development.

Recent studies confirm an instructive action
cytokines

A number of studies have now shown that cytokines have an instructive action on cell-fate decisions. M-CSF, G-CSF, and erythropoietin (Epo) instruct monocyte, neutrophilic and erythroid fates, respectively. In 2009, Rieger et al., by monitoring individual hematopoietic progenitors in culture, confirmed that M-CSF and G-CSF provided instructive cues. Using bioluminescence techniques and a LysM-GFP reporter system to detect differentiation and cell death, the group observed that individual GMPs adopted monocytic or neutrophilic fates in the presence of M-CSF or G-CSF, respectively26. More recently, Mossadegh-Keller et al. made use of PU.1-GFP reporter mice to show that M-CSF drives expression of myeloid-associated genes in some LT-HSCs. In vivo, M-CSF stimulated expression of PU.1, HSCs generated increased numbers of GMPs in the spleen and peripheral myeloid cells at the expense of cells undergoing megakaryocyte, erythroid, and lymphoid development when compared with non-primed Pu.1- HSCs27. Epo induces priming of erythroid lineage-associated genes in LT-HSCs and in vivo skew the potential of these cells towards an erythroid fate14. An increase in serum Epo levels in mice led to the expansion of committed erythroid and megakaryocyte and erythroid progenitors (MEP) in the bone marrow, whereas megakaryocyte progenitors, pre-GMP, and LMPP populations were decreased.

An instructive action of Flt3 ligand in determining
cell fate

One of the cytokines that is essential to cell survival and proliferation during early hematopoiesis is the ligand for Flt350,51. Flt3 ligand (Flt3L) was described two decades ago26 and is the only known ligand for Flt3. Upon ligand binding, the Flt3 receptor dimerizes and initiates signaling that involve STAT5a, ERK1/2, and PI3K12. Flt3 is an important area of research since mutations in Flt3 were among the first ones discovered in acute myeloid leukemia (AML)18. Presently, there is substantial interest in elucidating the instructive role of Flt3L.

Flt3 expression occurs during hematopoiesis at the non-self-renewing ST-HSC stage of development51. In fact, Flt3 up-regulation relates to the loss of self-renewal capacity69. This may just be a coincidence, or Flt3/Flt3L provokes the loss of self-renewal capacity by an as-yet-unknown mechanism. Thereafter, MMP express Flt3 as do several downstream progenitors with myeloid and/or lymphoid potential, while the MEP is Flt351. As lineage options become more restricted, Flt3 expression is down-regulated with the exception of DCs60. Importantly, Flt3L exerts a role by interacting with other cytokines such as IL-7 or SCF62,63. For example, IL-7 and

![Figure 5](image_url)  
In Figure 5, instructive action of Flt3 ligand in determining lymphoid/myeloid versus MegE lineage development. If the Flt3 ligand signal strength exceeds a certain threshold level, cells enter the lymphoid/myeloid branch at the expense of the MegE lineage. HSC, hematopoietic stem cell; MPP, multi-potent progenitor.

Flt3L stimulate lymphoid development in a coordinated manner that occurs in a narrow window during which cells express both receptors. In the case of in vitro cultures of hematopoietic progenitor cells, such as FETP, CLP, or EPLM, and in the presence of IL-7, Flt3L provides an additive anti-apoptotic effect while stimulating proliferation64,65 (and our observations). This additive effect is reported to result from parallel activation of the IL-7 receptor (IL-7R) and Flt3 via separate signaling pathways converging to activate Stat565.

Mice with targeted gene disruption of Flt366 or its ligand67 have provided precise information about the actions of Flt3/ Flt3L. These mice have defects in the developmental potential of myeloid/lymphoid progenitors as well as reduced numbers of B cells, DCs, and NK cells66,67. Tsapogas et al. have provided evidence to support the notion that Flt3L is instructive to cell decision-making. These workers generated a Flt3L transgenic (Flt3L-Tg) mouse that expresses human Flt3L68. Flt3L-Tg mice have a tremendous expansion of hematopoietic progenitors in the bone marrow. Of all the progenitor populations analyzed, the only progenitor that was decreased was the MEP. As such, the Flt3-Tg mice had decreased platelet counts and developed anemia. Previous studies had reported that the expression of Flt3 after the ST-HSC stage leads to a reduction of megakaryocyte and erythroid potentials21. Also, a significant reduction in MEP progenitors was observed by day 3 when wild-type mice were injected with recombinant Flt3L59. Considering this rapid response, it is likely that the reduction in MEP numbers was a consequence of a Flt3L threshold response in upstream Flt3+ progenitors rather than a space restriction within the bone marrow caused by the over-proliferation of other progenitors. It was proposed that upstream Flt3+ progenitors develop towards lymphoid/myeloid lineages on receiving Flt3L stimulation above a certain threshold level and develop towards the MegE (megakaryocyte/erythroid) lineage if this threshold level was not reached. In other words, an increased level of Flt3L guides the development of cells towards the lymphoid/myeloid fates at the expense of the MegE fates (Figure 5). This provides an explanation of the MegE developmental defect in the Flt3L-Tg mice. The exact mechanism by which Flt3L exerts an instructive action remains to be elucidated and whether this occurs at the CLP/EPLM level is of considerable interest.
Cytokines play different roles at various developmental stages

This is best illustrated by consideration of the actions of IL-7 at various stages of B and T lymphocyte development\(^{60-72}\). These include promoting cell survival and proliferation and facilitating decision-making during differentiation (Figure 6). IL-7 was described in 1988\(^{73}\). In mice, deficiencies in IL-7 or the receptor lead to impairment of both B and T lymphopoesis\(^{44-56}\). IL-7 is not required for B lymphopoiesis in humans, as deficiencies in IL-7 result in apparently normal B cell development\(^7\). However, the B cells could be the result of fetal development since they can be cultured in vitro in the absence of IL-7 while postnatal B cells seem to be IL-7 dependent\(^73\).

CLPs express IL-7R, which provides a criterion used to isolate these cells\(^1\). CLPs are dramatically reduced in IL-7-deficient mice\(^8\), indicating that this cell is critically dependent on IL-7. EPLMs also express IL-7R\(^9\). A strong proliferative action of IL-7 is seen when CLPs and EPLMs, sorted from mouse bone marrow, are cultured on OP9 stromal cells and treated with IL-7\(^3\) (and our own observations). Furthermore, these cells are able to differentiate to the next B cell developmental stage, the pro-B (pre-B1), indicating that IL-7 acts as a differentiation factor at the CLP/EPLM level. In keeping with this notion is that CLP cells are unable to reach the pro-B (pre-B1) stage in \(\gamma c^-\), IL-7R\(\alpha^-\), or IL-7\(^-/-\) mice, and instead arrest at an uncommitted level\(^73\). These data suggest an instructive role for IL-7. This appears to occur via IL-7-induced STAT5 signaling which regulates the expression levels of EBF (early B cell factor)\(^30,80\). In turn, EBF activates transcription of Pax5\(^81,82\) leading to the expression of CDO1. This cell surface marker indicates that cells have gained the pro-B phenotype and commitment to the B cell lineage\(^83,84\).

The pro-B (pre-B1) stage is also sensitive to the action of IL-7. Fetal and bone marrow pro-B (pre-B1) cells can be cultured in vitro in the presence of IL-7 and stromal cells for long periods of time (more than 4 months)\(^85\). The survival action of IL-7 appears to be exerted via expression of the gene encoding the anti-apoptotic protein Mcl-1\(^86\). Moreover, there is evidence to suggest that IL-7 facilitates differentiation of pro-B (pre-B1) cells into cytoplasmic \(\mu\) immunoglobulin (IgM) heavy chain expressing large pre-B2 cells\(^87,88\) (and our own unpublished observation). IL-7 is not instructive, since cells differentiate irrespective of the presence of IL-7\(^87,88\). Instead, the data argue for a permissive role of IL-7 in promoting survival and proliferation during the pro-B (pre-B1)/large pre-B2 transition. However, some groups have reported that STAT5-mediated IL-7 signaling controls chromatin accessibility and the rearrangement of distal VH genes at the \(\lg h\) locus\(^89-91\). This is exemplified by a significant decrease in distal VH rearrangements in B220\(^7\) IgM\(^+\) bone marrow B lymphocytes of IL-7\(^-/-\) and Stat5\(^-/-\)-deficient mice\(^90,91\). In contrast, Malin has argued that there is no substantial difference in the distal VH rearrangement genes seen for STAT5 deficient, IL-7R mutant, and control pro-B (pre-B1) cells\(^89\). Finally, as to the pre-B1/pre-B2 transition, it has been reported that IL-7 acts to prevent premature rearrangements of the Igκ via binding of IL-7-mediated STAT5 to the Igκ intrinsic enhancer (IκE\(s\))\(^92\).

The large pre-B2 transition is the last stage in B cell development that is sensitive to the action of IL-7. The proliferation of these cells is improved by the presence of IL-7\(^93\). However, IL-7 acts as an anti-differentiation factor towards immature B cells by blocking the rearrangement of the light chain loci\(^94,95\). Upon withdrawal of IL-7 from the culture, cells undergo Igκ recombination and differentiation into IgM positive cells\(^96,97,98\). At this stage, as a consequence of the pre-B cell receptor signaling, IL-7R is
down-regulated and the later stages of B cell development are unresponsive to IL-7Rα9,10,29.

During T lymphocyte development, the thymus is seeded by ETPs32 and there is a general agreement that these cells can give rise to multiple lineages. In humans, the ETP retain lymphoid/myeloid potentials and can generate separate lymphoid- and myeloid-committed progenitors. Lymphoid-restricted progenitors seem to be the main thymus colonizers in the mouse33, although several in vitro studies suggest these cells have myeloid potentials8,9. IL-7Rα is a direct target of Notch1101, the master signaling pathway that regulates thymopoiesis73-76. Up-regulation of IL-7Rα at the ETP stage100 is a hallmark for lymphoid commitment, and IL-7-mediated signaling triggers the first wave of expansion of lymphoid-committed progenitors98,101. IL-7 is essential to survival and proliferation of ETP, and appears to transmit signals for the survival and proliferation of lymphoid-committed cells at the expense of the myeloid branch75. Sustained Notch signaling favors cell development along the T cell lineage by inducing the transcription of important T lymphocyte differentiation factors and blocking the development of cells towards other lineages105-108. IL-7 is also required for the transition of ETP cells to the DN2 stage in vitro64. Moreover, an early block is observed in T lymphocyte development in IL-7 or IL-7R-deficient mice99. However, Bcl-2 is sufficient to rescue T lymphocyte development (at least the α/β branch of T lymphocytes), thus confirming a permissive role for IL-7 during the early thymopoiesis109.

IL-7Rα chain expression increases progressively until the DN2 stage, coinciding with the first massive cellular expansion, and then steadily decreases. DN2 cells with high IL-7R expression levels are diverted to the γδ T cell lineage, at least in mice111, and IL-7R signaling controls accessibility to the T cell receptor (TCR)/γ locus and its rearrangement112,113. Thus, IL-7 plays an instructive role in this developmental process and Bcl-2 is unable to rescue the development of γδ T cells114,115. The cells with reduced IL-7R expression and/or limited IL-7 availability progress to the DN3 stage. The proliferation of these cells is still IL-7 dependent104,105,109. However, a diminished IL-7R signaling seems necessary for DN2 mouse thymocytes to up-regulate Bcl11b, a transcription factor that is essential to the T cell lineage116. Thereafter, on successful rearrangement of the TCR α chain and subsequent pre-TCR expression, IL-7-R is down-regulated, and cells move to the DN4 stage105,107,109. Here, at the β selection checkpoint, there is a second wave of expansion that is controlled by Notch1 signaling and pre-TCR expression64. As such, the DN3 to DP (double-positive) transition is IL-7 independent. However, the continued presence of IL-7 in in vitro cultures blocks differentiation and DP cells are generated only on IL-7 removal. In this case, IL-7 is acting as an anti-differentiation factor64. This highlights the importance of the IL-7-R signaling suppression during the DP transition, which is guaranteed by SOCS-1 (suppressor of cytokine signaling)117. After the DP selection stage, IL-7R surface expression levels are restored and IL-7 signaling seems to be required for the DP to single positive CD8 transition118. In keeping with this, Park et al. reported differentiation of single positive CD8 cells (CD8+) and RUNX3 transcription factor up-regulation in transgenic TCR mice expressing a transgene-derived IL-7-Rα chain in an IL-7 over-expressed environment119. Finally, IL-7 is required for homeostatic expansion of naïve CD8+ and CD4+ T cells120,121.

In the context of the role of IL-7 in lymphopoiesis and in particular the flexibility of the hematopoietic process, the effect of pregnancy on hematopoiesis and lymphopoiesis is worthy of mention. Pregnancy is a physiological process in which there is an increase in blood volume and the bone marrow responds by increasing erythropoiesis122. Lymphopoiesis is also affected, and two mechanisms account for the reduced B lymphopoiesis observed in pregnancy. First, early B progenitor cells are sensitive to the increased levels of sex steroids in pregnancy122, and second, in mice, IL-7 production and availability dramatically decrease123. Not only is the bone marrow affected during pregnancy but also the maternal thymus undergoes dramatic involution124. Both B and T lymphopoieses return to pre-pregnancy levels following parturition and weaning. A full explanation of these associated phenomena is not currently available and is certainly worthy of further investigation. In regard to the effect of cytokines on lymphopoiesis in the developing fetus, an interesting and recent finding is that inflammatory cytokines interferon-α (IFN-α) and IFN-γ positively regulate the number of lymphoid progenitors in mouse embryos, and IFN-γ signaling also affects the number of HSCs122. These findings have important implications to enhancement of the formation of hematopoietic stem cell and progenitor cells in the embryo and the differential regulation of hematopoiesis during fetal and adult life.

Implications of heterogeneity and versatility to leukemia

A Darwinian viewpoint on leukemia

A widely-held viewpoint for a number of years is that the cell population that makes up an overt leukemia and solid tumor is heterogeneous and that there is a hierarchical organization. A small subset of cells, which are at the apex of the hierarchy, is responsible for maintaining the long-term growth of the leukemia and tumor mass. These leukemia- and cancer-maintaining cells, which have an inherent capacity to self-renew, are termed leukemia stem cells (LSC) and cancer stem cells (CSC), respectively126,127. There is also the question of the normal cell from which the leukemia/tumor has arisen, which are termed leukemia-initiating cells (LIC) and cancer-initiating cells (CIC), respectively. These are the cells that suffered an oncogenic lesion and that evolve to give rise to the leukemia/tumor. And, many leukemias and cancers arise from a self-renewing stem cell or downstream progenitor if self-renewal is reactivated. The origin and LSC/CSC nature of leukemias/cancers is important to putting into practice a cancer stem cell-based therapeutic to cure patients128-132.

To add to the difficulty of designing therapies to eliminate LSC/CSC, Greaves has proposed a Darwinian viewpoint on the nature of CSC which pays attention to the dynamics of the cancer133,134. Studies of single acute lymphoblastic leukemia (ALL) cells have revealed that the leukemia "stem" cells are genetically diverse. Greaves has likened the complex and branching clonal architecture of ALL "stem" cells to Darwin's evolutionary tree-like divergence diagram which was drawn in 1837. To add to the Darwinian analogy, as the leukemia evolves,
in an almost entirely clinically silent manner, cells acquire gene copy number alterations. The cytokine TGF-β has been proposed to exert a selective Darwinian advantage to expand the cells that are first “fit” as to genomic alteration and at risk of acquiring secondary multiple gene alterations. In essence, this is similar to “natural selection” as described by Darwin. Often what we discover from studies of cancer cells, and ascribe to these cells, turns out to be a feature of normal cells. A simple example is that the common ALL-associated antigen (CD10) was first described as a candidate leukemia-specific antigen. In fact, CD10 is expressed by rare B cell progenitor cells and its presence on ALL cells was telling us something about the origin of common ALL. So, we might view normal hematopoiesis as a Darwinian process that is driven by selective pressures, namely cytokines, acting on an inherent diversity that is sufficient to generate the various types of blood and immune cells. In essence, inherent cell heterogeneity must have been the template to the evolution of the wide variety of immune cells that exist in higher mammals.

Are leukemia stem cells as versatile as their normal counterparts?

An important question that arises from the above considerations is the following: are LSCs as versatile as normal HSCs in terms of their capacity to access routes to end-cell types? There is evidence to suggest that LSCs are less versatile in this regard, as exemplified by erythroleukemia [acute myeloid leukemia (AML) FAB-M6] and pre-B/pro-B common acute lymphoblastic leukemia (ALL). The cells that sustain these leukemias appear to have become directed to generate cells of a particular cell type. An accumulation of erythroid precursors and myeloblasts in AML FAB-M6 reveals a disease origin in a cell with multi-lineage potential. However, the partial differentiation of the leukemic blast cells is restricted to certain pathways, and accordingly disease sub-sets are characterized as myeloblast-rich (FAB-M6A), proerythroblast-rich (FAB-M6B) and myeloblast- and proerythroblast-rich (FAB-M6C). In the case of childhood ALL, including pre-B ALL, pro-B ALL, and common ALL (c-ALL), there are arguments to support a disease origin in either a cell that is committed to B lymphocyte development or a cell that is more stem cell-like. The former notion is a long-held viewpoint. In favor of the latter cellular origin is the c-ALL-derived cells lacking the B-lineage markers CD10 and CD19 and expressing the stem cell marker CD34 can give rise to c-ALL and pre-B ALL when transplanted into mice. The argument about the precise “target” cell that is transformed in childhood ALL could be set aside as versatility of lineage options extends to progenitor cells, as illustrated by the cytokine-mediated redirection of CLPs to the myeloid lineage. Strikingly, the blast cells that accumulate in the blood and bone marrow in childhood ALL are restricted to B-lineage development.

Conclusions and outlook

There is good evidence to support the notions of developmental plasticity in the hematopoietic system and that growth factors can instruct lineage potentials. This leads to interesting questions as to how HSCs establish and, as required for steady-state hematopoiesis, maintain pluripotency, and how a network of factors might participate to control HSC identity and commitment ability to drive contributions to homeostasis and adaption to inflammatory conditions. Presumably to maintain HSC pluripotency, there is the requirement to control genomic stability, in part via controlling intrinsic DNA repair machinery. As mentioned above, vWF+ HSCs require Tpo for their maintenance. This growth factor also plays a role in regulating HSC genomic stability, since the efficiency of DNA-Protein Kinase-dependent DNA repair, in response to DNA damage, is increased by Tpo. Tpo-induced activation of ERK and NF-xB in HSCs is important to damage repair. As to Tpo playing a role to ensure chromosomal integrity, the plot thickens regarding the roles of growth factors. Ineffective DNA repair of genomic instability in LIC and/or LSC might curtail the availability of lineage options of these cells. Global genomic instability is unusual for AML, and for both AML and pre-B ALL, a very small number of mutations are required to generate the leukemic clone. Genomic instability is a feature of the chronic phase of chronic myeloid leukemia, resulting in RCR-ABL-1 mutations that encode resistance to the tyrosine kinase inhibitors (for example imatinib) used to treat the disease. Importantly, instability has been postulated to occur in primitive leukemia progenitor cells in patients who have not been treated with tyrosine kinase inhibitors. Also, it has been suggested that the instability is due to high levels of DNA damage by reactive oxygen species and inefficient/faithful repair of DNA double-strand breaks that lead to chromosomal aberrations.

In summary, for normal stem and progenitor cells, the versatile nature of the pair-wise model provides an appropriate template for the persuasive action of growth factors to shape hematopoiesis. As yet, we do not know whether LIC and/or LSC are as or less versatile in terms of lineage options than their normal counterparts. Selective instability in genomic elements encoding controls on lineage decision-making (for example transcription factors and signaling molecules) might restrict the versatility of LIC/LSC.

Declaration of interest

The authors alone are responsible for the content and writing of the paper.

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013 under Research Executive Agency grant agreement No 315902. C.J.M., L.A.S. and L.v.M. gratefully acknowledge receipt of a Marie Curie Research Associate post, G.B., K.T., R.C. and A.R. are partners within the Marie Curie Initial Training Network DECIDE (Decision-making within cells and differentiation entitcessies entiett).


36. Guinan C, Francis JR, Bed DJ, et al. Differentiation restriction in the stemcell granulocyte, macrophage, eosinophil-granulocyte-


49. Konide M, Schuerz DC, Miyamoto T, et al. Cell fate conversion of lymphoid-committed progenitors by instructive actions of cyto-


10. Acknowledgements

I want to thank all people who supported me during my time as a PhD student.

My first thanks go to Ton for allowing me to take part of your brilliant team, it has been a great honour! Thank you for your guidance and sharing your enormous expertise and knowledge as well as for maintaining always with a laugh the good environment characteristic of the lab. Moreover, I want to thank you for encouraging me to explore a new scientific dimension, the bioinformatics and for providing me the tools, time and comprehension to learn and grow in this field. In summary, thanks for your always support!

Secondly I would like to thank Daniela and Rhodri, who kindly agreed to become members of my PhD committee. I received great input and feedback during numerous lab meetings as well as DECIDE meetings. Thank you, Rhodri, for your time in Basel, for our discussions and for your enormously helpful corrections on the thesis.

The third, but not less important, big thank you goes to all current and former members of “Team Ton”. It has been a great team with an enviable atmosphere and a helping hand whenever needed as well as scientific advice. Thank you Fabian, Lilly, Panos, Patrick, Stefan, Marianne, Corinne, Mike, Giusy, Jonas, Hannie, Nicole, Simon, Ilda, Sofia and Jan. Specially, I want to thank Lilly, who has been also part of the DECIDE network, for all the moments and support to each other; Panos, for all the discussions about science and the world in general, for your scientific input, encouraging words and for the time dedicated to fine tune the thesis; Patrick, for all the fun and for your enormous and contagious positivism and enthusiasm in the lab, you have been a great support! Moreover, I also want to thank Daniela’s and Roxanne’s group for sharing the lab seminars and helpful discussions. Besides the lab work I enjoyed all the fun we had at Christmas parties, lab outings, coffee breaks, late beers in the lab, retreats, the use of the hand-made “signs” etc
I would also like to thank the DECIDE network. I feel very proud and fortunate to have taken part in this PhD Marie Curie programme. I really want to thank all the people that have made this network possible because it has been a huge opportunity for us. It was great to bring all the fellows, supervisors and partners of the DECIDE network together once a year in a different country. Thank you to all the fellows, I am so happy to have met you and share our parallel PhD experiences at different countries. Let’s keep in touch with each other!

Finally, to all the people outside science who has supported me. My girls, Flavia, Marina and Marian, for your valuable distance friendship, understanding and encouraging me on keeping going during my PhD. Also for our yearly holidays where I could refill the energy and keep the happiness and motivation. Thanks to all the people who have contributed to my happy stay in Basel: Basilenyus, DBM community, Marwa, Lina, Maria etc. but above all, to Fran. Cariño, no tengo palabras para expresarte el agradecimiento que siento hacia ti. Sin lugar a duda has sido la persona que has vivido más de cerca el inicio, desarrollo y culminación de mi tesis doctoral. Tu apoyo ha sido fundamental, siempre has tenido palabras de ánimo y comprensión (así como hechos). Vine a Basel para embarcarme en el proyecto de hacer un doctorado, pero gracias a ti esta experiencia se ha convertido en algo inolvidable, juntos hemos descubierto un nuevo país, nos hemos adaptado y hemos tirado hacia delante. Eres esa persona esencial que necesito en mi vida para futuros proyectos y a la que también pienso apoyar en todos los que tengas. Gracias de corazón por ser ante todo mi amigo y confidente, el mejor compañero de viaje. Te quiero.

També vull agrair a tots els amics, sobretot de Mallorca, per entendre lo enfeinada que he estat escrivint. No vull acabar sense agrair des de lo més profund del cor a tota la meva família i, sobretot en es meus pares i germà. Papas moltes gràcils per haver confiat sempre amb jo i per haver facilitat sempre que compleixi es meus somnis. No seria on sóc si no fos per voltros. Per molt lluny que estiguem, sempre vos he sentit a prop. A més, aquests darrers dies escrivint des de casa (Mallorca) m’han donat energies extres per acabar sa tesi. Ha estat molt d’esforç però ha valgut la pena. Moltes gràcils per tot, vos estim.