

Discovery of a CD10 negative B-progenitor in human fetal life identifies unique ontogeny-related developmental programs

Short running title: Fetal specific human B-cell developmental programs

Sorcha O'Byrne¹, Natalina Elliott¹, Siobhan Rice², Gemma Buck^{1,2}, **Nicholas Fordham², Catherine Garnett²**, Laura Godfrey², Nicholas T. Crump², Gary Wright³, Sarah Inglott³, **Peng Hua^{2,4}, Bethan Psaila²**, Benjamin Povinelli², **David J.H.F. Knapp²**, Antonio Agraz-Doblas^{5,6}, Clara Bueno⁶, Ignacio Varela⁵, Phillip Bennett⁷, Hashem Koochy⁸, **Suzanne M Watt⁴**, Anastasios Karadimitris⁹, Adam Mead^{2,10}, Phillip Ancliff³, Paresh Vyas^{2,10}, Pablo Menendez^{6, 11, 12}, Thomas. A. Milne^{2,10}, Irene Roberts*^{1,2,10}, Anindita Roy*¹

¹Department of Paediatrics, University of Oxford, Oxford, UK

²Medical Research Council (MRC) Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK.

³Department of Haematology, Great Ormond Street Hospital NHS Foundation Trust, London, UK.

⁴Stem Cell Research, Nuffield Division of Clinical Laboratory Sciences, Radcliffe Department of Medicine, University of Oxford, Oxford, UK.

⁵Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), Universidad de Cantabria-CSIC, Santander, Spain.

⁶Josep Carreras Leukemia Research Institute-Campus Clinic, Department of Biomedicine, School of Medicine, University of Barcelona, Barcelona, Spain.

⁷Institute of Reproductive and Developmental Biology, Department of Surgery & Cancer, Hammersmith Hospital campus, Imperial College London, UK.

⁸MRC Human Immunology Unit, MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK.

⁹Centre for Haematology, Department of Medicine, Hammersmith Hospital, Imperial College London, London, UK.

¹⁰Haematology Theme, Oxford Biomedical Research Centre, Oxford University Hospitals, Oxford, UK.

¹¹Institucio Catalana of Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

¹²Centro de Investigación Biomédica en Red en cancer (CIBERONC)-ISCIII, Barcelona, Spain.

*These authors contributed equally.

Correspondence:

Dr Anindita Roy

University of Oxford Department of Paediatrics

Level 2, Children's Hospital, John Radcliffe

Oxford OX3 9DU

Phone: +44(0)1865 222 419

anindita.roy@paediatrics.ox.ac.uk or irene.roberts@paediatrics.ox.ac.uk

Text word count: 3999 (4000)

Abstract word count: 248

Number of figures and tables: 5 (+11 supplemental figures and 9 supplemental tables)

Number of references: 56

KEY POINTS

- CD10-ve PreProB-progenitors are the earliest fetal B-lymphoid restricted progenitors and are enriched in fetal bone marrow
- Fetal PreProB-progenitors have a unique ontogeny-related developmental gene expression program distinct from their rare adult counterparts

ABSTRACT (247 words)

Human lymphopoiesis is a dynamic life-long process that starts *in utero* 6 weeks **post-conception**. Fetal B-lymphopoiesis remains poorly defined and yet is key to understanding leukemia initiation in early life. Here, we provide a comprehensive analysis of the human fetal B-cell developmental hierarchy. We report the presence in fetal tissues of two distinct CD19+ B-progenitors, an adult-type CD10+ve ProB-progenitor and a new CD10-ve PreProB-progenitor, and describe their molecular and functional characteristics. PreProB- and ProB-progenitors appear early in the first trimester in embryonic liver, followed by a sustained second wave of B-progenitor development in fetal BM, where together they form >40% of the total HSC/progenitor pool. Almost one-third of fetal B-progenitors are CD10-ve PreProB-progenitors while, by contrast, PreProB-progenitors are almost undetectable ($0.53\pm0.24\%$) in adult BM. Single-cell transcriptomics and functional assays place fetal PreProB-upstream of ProB-progenitors, identifying them as the first B-lymphoid restricted progenitor in human fetal life. Fetal BM PreProB- and ProB-progenitors both give rise solely to B-lineage cells yet they are transcriptionally distinct. Like their fetal counterparts, adult BM PreProB-progenitors give rise only to B-lineage cells *in vitro* and express the expected B-lineage gene expression program. However, fetal PreProB-progenitors, display a distinct, ontogeny-related gene expression pattern which is not seen in adult PreProB-progenitors; and share transcriptomic signatures with CD10-ve B-progenitor infant acute lymphoblastic leukemia blast cells. These data identify PreProB-progenitors as the earliest B-lymphoid-restricted progenitor in human fetal life, and suggest that this fetal-restricted committed B-progenitor might provide a permissive cellular context for prenatal B-progenitor leukemia initiation.

INTRODUCTION

The developmental hierarchy of human fetal B-lymphopoiesis has not been fully defined. The presence of B-lineage cells has been noted in human fetal liver (FL) by 7-8 **post-conceptual** weeks (**pcw**) and in fetal blood and bone marrow (BM) by 12 **pcw**¹⁻⁶, but their functional and molecular properties were not characterized. Indeed, it is only recently that progenitors with B-lineage potential present exclusively or principally in fetal rather than adult tissues have been reported in humans. **Interest in these cells is considerable as potential targets for leukemic transformation given that many childhood and all infant B-lineage acute lymphoblastic leukemia (ALL) originate in fetal life**⁷. A CD19-CD127+ progenitor with multi-lymphoid or lympho-myeloid potential has been described in FL^{8,9} and a CD7+ progenitor with lympho-myeloid output in fetal BM¹⁰. However, expression of CD19 is considered the definitive marker of B-cell commitment and in adult BM the earliest B-lineage restricted progenitor reported is a ProB-progenitor (CD34+CD10+CD19+) (reviewed in¹¹). The presence of a second B-progenitor, a PreProB-progenitor (CD34+CD10-CD19+) has been noted in cord blood^{12,13} and while we also noted a CD34+CD10-CD19+ population in second trimester FL, we did not characterize these cells at the molecular or functional level¹⁴. Thus, **the identity of the earliest committed B-progenitor in human fetal life remains unclear.**

To address this, we carried out comprehensive immunophenotypic, functional and molecular characterization of human fetal B-lymphoid development to define the fetal B-cell developmental hierarchy for the first time. We show that by **7 pcw** both CD10-ve PreProB- and CD10+ve ProB-progenitors are present in **embryonic liver** before a second wave of B-progenitor development is initiated in fetal BM 4 weeks later. We demonstrate for the first time that in fetal tissues, PreProB-progenitors are the earliest B-lymphoid restricted progenitors in the B-cell developmental hierarchy and that they lie upstream of ProB-progenitors from which they are transcriptionally distinct, and downstream of early lymphoid progenitors (ELP; **Lin**-CD34+CD127+CD10-CD19-).

METHODS

Samples: Fetal tissue was collected after written informed consent and approved by the Hammersmith, Queen Charlotte's & Chelsea and Acton Hospitals Ethics Committee (04/Q0406/145). Donated fetal tissue was also provided by the **Human Developmental Biology Resource (HDBR, www.hdbr.org)** regulated by the UK Human Tissue Authority (HTA, www.hta.gov.uk). Normal adult BM was collected during elective orthopedic surgery after written informed consent and approved by INForMeD HRA approval (IRAS 199833; REC no. 16/LO/1376). Otherwise, adult BM MNC were purchased from StemCell Technologies, Canada and Lonza, USA. Infant ALL samples used for comparison are from published datasets^{15,16}.

Flow cytometric analysis and sorting, including cell cycle analysis: See Supplemental methods.

MS-5 stromal co-cultures: 100 flow-sorted fetal BM CD34+ HSPC were co-cultured on confluent MS-5 stroma in α MEM (Gibco) supplemented with 10% heat-inactivated batch-tested FBS, 100U/mL Penicillin, 100 μ g/mL Streptomycin, 2mM L-glutamine, 50 μ M 2-Mercaptoethanol, 10mM HEPES, SCF (20ng/mL), Flt3L (10ng/mL), IL-2 (10ng/mL) and IL-7 (5ng/mL) in 24-well plates. Cultures were maintained and output analyzed as previously described¹⁴.

OP9-DL1 stromal co-culture: OP9 stromal cells were a kind gift from Prof Adrian Thrasher (Institute of Child Health, University College, London). OP9-DL1 stromal cells were grown to confluency and passaged as described previously¹⁷. See Supplemental methods for further details.

Fetal MSC stromal co-cultures: Fetal BM MSC were isolated by plating $1-2 \times 10^6$ freshly-isolated MNC in α MEM supplemented with 10% heat-inactivated FBS, 100U/mL Penicillin, 100 μ g/mL Streptomycin, 2mM L-glutamine, 50 μ M 2-Mercaptoethanol and 10mM HEPES. After 72-96hrs, non-adherent cells were removed and adherent cells cultured until confluent before being passaged using 0.05% trypsin/EDTA. MSC were immunophenotypically characterized by flow cytometry¹⁸. 24 hours prior to co-culture MSC were irradiated with 20Gy and plated in 24-well plates at 0.05×10^6 - 0.1×10^6 cells/well to

reach confluence within 24hrs. 500-1000 sorted fetal BM HSPC were seeded/well and co-cultured with cytokines as described for MS-5 co-cultures. Co-cultures were replenished with half-media changes every 3-4 days. Wells were harvested at days 3, 7, 14, 21 and 28 for analysis by flow cytometry.

Clonogenic assays: Myeloid and erythroid differentiation potential of 100 sorted HSPC were assessed using colony-forming assays as previously described¹⁴.

Bulk RNA-Sequencing: For fetal BM bulk RNA-Sequencing, the SMART-Seq2 protocol was followed¹⁹. Briefly, 100 purified HSPC were sorted into lysis buffer containing 0.4% Triton X-100 (Sigma-Aldrich), RNase inhibitor (Clontech), 2.5 mM dNTPs (Thermo Fisher) and 2.5 μ M oligo-dT30VN primer (Biomers.net). cDNA was generated using SuperScript II (Invitrogen), pre-amplified using KAPA HiFi HotStart ReadyMix (KAPA Biosystems) using 18 cycles of amplification. After PCR amplification, cDNA libraries were purified with AMPure XP beads (Beckman Coulter) according to manufacturer's instructions. Post-purification libraries were resuspended in EB buffer (Qiagen). Quality of cDNA traces was assessed using a High Sensitivity DNA Kit in a Bioanalyzer instrument (Agilent Technologies). Tagmentation and library preparation were performed using the Nextera XT DNA Library Preparation Kit (Illumina) according to manufacturer's instructions. Libraries were sequenced on an Illumina HiSeq4000 to generate 75bp paired-end reads, yielding >30 million reads/sample.

Single cell qRT-PCR: Gene expression was assessed using the Biomark HD microfluidics system (Fluidigm). Single HSPC were index-sorted into 96-well plates containing pre-amplification mix and processed as previously described²⁰. Sorted cells were simultaneously analyzed for relative levels of expression of genes with known roles in **hematopoiesis**, B-lymphopoiesis and/or leukemia (**supplemental Tables 2 and 3**). Gene expression was normalized to average expression of three housekeeping genes (B2M, GAPDH, ACTB) **and plotted using ggplot2**. Diffusion pseudotime analysis was performed on log₂(relative expression + 1) data using the R package 'destiny' (version 2.10.2) with the distance metric set to "rankcor"^{21,22}. **Scatter plots were generated using Partek software. Copyright, Partek Inc, St Louis, MO, USA.**

IgH rearrangement analysis: Samples were screened for IgH complete (VH-DH-JH) and IgH incomplete (DH-JH) rearrangements using BIOMED-2 protocols to detect clonality. Cells were FACS-sorted into PBS and the pellets frozen. DNA was whole genome amplified using a Qiagen REPLI-g single cell kit. Heteroduplex analysis was used after PCR reactions²³. PCR products were excized and eluted from polyacrylamide gels and sequenced using Thermo Fisher Scientific BigDye® Terminator v3.1 Cycle Sequencing Kit protocol. Capillary electrophoresis was performed on a Thermo Fisher Scientific 24-capillary 3500xL Genetic Analyzer. Sequences were entered into National Center for Biotechnology Information IgBLAST tool for immunoglobulin domain identification and alignment and annotated using IMGT/V-QUEST²⁴.

ATAC-sequencing: The protocol was adapted for small cell numbers from Corces *et al*²⁵. 2000 PreProB- and ProB-progenitors (n=3); and 5000 Lin+ (CD2+CD3+CD14+CD16+CD56+CD235a+) MNC were sorted into 25µL transposase mix (Tn5 enzyme, 2X TD buffer and 0.01% digitonin, Promega) at 4°C then immediately incubated at 37°C for 30 minutes at 300 rpm. Transposed DNA was purified using a Qiagen MinElute kit and eluted in 10µL pre-warmed EB. Transposed fragments were amplified by PCR as previously described²⁶. Libraries were quantified by qPCR (NEBNext Library Quant for Illumina kit) before sequencing using a 75-cycle paired-end kit on the NextSeq Illumina platform.

Statistics

T-tests, Mann Whitney tests, Wilcoxon test and ANOVA followed by multiple comparisons testing were used to compare experimental groups as indicated in the figure legends. Statistical analyses were performed using GraphPad Prism v7.00 or R v3.4.1. Data are expressed as mean ± SEM unless otherwise indicated.

Data availability

Fetal BM RNA-Sequencing and ATAC-sequencing files have been deposited in NCBI's Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/query>) with accession numbers GSE122982 and GSE122989 respectively. For all other original data, please contact anindita.roy@paediatrics.ox.ac.uk.

167

168

169

RESULTS

Onset and expansion of human B-lymphopoiesis occurs sequentially in fetal liver and fetal BM

To determine the onset of B-lymphopoiesis in fetal life we measured the frequency of B-cells and their progenitors from 6-21 pcw using the immunophenotypic criteria shown in supplemental Table 4. In FL, we found that immunophenotypic hematopoietic stem cells (HSC), multipotent progenitors (MPP) and a low frequency of lymphoid-primed multipotent progenitors (LMPP) and ELP are already present at the embryonic stage by 6 pcw (Figure 1a-c), while PreProB- and ProB-progenitors, and small numbers of B-cells (CD34-CD19+) become detectable a week later (Figure 1d-f).

By contrast, we detected no CD34+ or B-cells in embryonic BM until after 10 pcw consistent with the absence of hematopoietic cells on 9-12 week fetal BM sections (data not shown) and as previously noted⁵. PreProB-, and ProB-progenitors (Figure 1g) are first seen in second trimester fetal BM at 11 pcw, appearing concurrently with immunophenotypic HSC/MPP, LMPP and ELP (Figure 1h). In FL, the frequency of PreProB- and ProB-progenitors remains low and constant ($2.5 \pm 0.4\%$ and $7.8 \pm 2\%$ respectively) through the second trimester. Meanwhile, in fetal BM, B-lymphopoiesis dramatically expands, with PreProB-progenitors peaking immediately at the onset of BM hematopoiesis ($19.4 \pm 6.1\%$ of CD34+ cells at 11 pcw) and ProB-progenitors rising steeply from $11 \pm 4.4\%$ at 11 pcw to $>30\%$ of all BM CD34+ cells towards the end of the second trimester (Figure 1d,e). Thus, PreProB- and ProB-progenitors are present at much higher frequency in the fetal BM at all gestations we assessed despite lower frequencies of HSC/MPP (Figure 1a,d,e,i). This also holds true when matched liver and BM from the same fetus are compared (Figure 1j, n=8). This contrasts with adult BM where PreProB-progenitors are virtually absent (Figure 2a). The high frequency of fetal BM PreProB- and ProB-progenitors in S-G2-M compared to HSC, MPP and LMPP (Figure 2b,c) suggests that B-progenitor proliferation underpins this rapid expansion rather than differentiation of upstream progenitors. Consistent with this, the frequency of CD34-CD19+ B-cells (most of which are CD10+ PreB cells²⁷) in fetal BM only starts to increase from 16 pcw (Figure 1h, Figure 2d).

These data show that although immunophenotypic PreProB- and ProB-progenitors appear first in embryonic liver at 7 pcw and persist there at low frequency in fetal life, the major

site of fetal B-lymphopoiesis is in fetal BM. A second wave of B-lymphopoiesis starting at 11 **pcw** fetal BM, is propagated by rapid expansion/self-renewal of PreProB-progenitors during a narrow time window in **early second** trimester followed by a switch to predominantly ProB-progenitor expansion and differentiation thereafter (Figure 2d).

Fetal BM PreProB-progenitors give rise solely to B-lineage cells in vitro and in vivo

To determine whether both PreProB- and ProB-progenitors are committed to the B-lineage or retain residual non-B lymphoid/myeloid output, we **first** cultured flow-sorted, immunophenotypically-defined fetal BM hematopoietic stem and progenitor cells (HSPC) (supplemental Figure 1) on MS-5 stroma (B/NK-cell and myeloid potential), OP9-DL1 stroma (T-cell potential) or in methylcellulose (myeloid/erythroid potential). PreProB- and ProB- progenitors both give rise to B-cells in MS-5 co-cultures with no T-cell (CD4+/CD8+) potential on OP9-DL1 co-cultures (Figure 3a,b; supplemental Figure 2a,b). By contrast, although ELP give rise to predominantly B-lineage cells, including PreProB- and ProB-progenitors and B-cells, they also generate CD34+CD10+CD19- progenitors, T-cells, NK-cells and a small number of myeloid cells, supporting their position upstream of committed B-progenitors (Figure 3a-c; supplemental Figure 2a,b). Fetal BM HSC, MPP and LMPP from the same samples also generate a mixture of myeloid, B-, NK- and T-cells (Figure 3a,b; supplemental Figure 2a,b). Similarly, in methylcellulose assays, myeloid clonogenic output was observed from HSC, MPP and LMPP, although not from ELP (Figure 3d). In addition, no evidence of engraftment of human cells (at 9 or 14 weeks) was seen after transplantation of PreProB- or ProB-progenitors (10,000 - 30,000 cells) into sub-lethally irradiated NSG (NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/SzJ) mice consistent with a lack of stem cell capacity (data not shown).

Next, to better capture the ability of PreProB- and ProB-progenitors to give rise to B-cells *in vivo* and assess any residual myeloid capacity, we measured engraftment of flow-sorted fetal BM PreProB-, ProB- and upstream Lin-CD34+CD19-CD10- progenitors 2 and 3 weeks after transplantation into NSG mice. In this short-term xenograft assay, PreProB-progenitors give rise to ProB-progenitors and B-cells only while upstream CD34+CD19-CD10- cells give rise to multilineage output, including ELP, PreProB- and ProB-progenitors (Figure 3e,f). ProB-progenitors give rise solely to B-cells although engraftment was very low (0.002-0.057% human CD45+ cells) and not accurately assessable in 2/4 cases. Together

these data show that PreProB-progenitors have no myeloid, T-cell or NK-cell output and, like ProB-progenitors, are committed to the B-lineage.

Finally, since MS-5 and OP9-DL1 co-cultures use murine stromal cells, we wondered if functional differences in the two types of fetal BM B-progenitor could be better identified using a more physiological microenvironment. To investigate this, and to determine the hierarchical relationship between PreProB- and ProB-progenitors, we co-cultured flow-sorted fetal BM PreProB- and ProB-progenitors on primary human fetal BM mesenchymal stromal cells (MSC) (supplemental Figure 3a) and measured the output of progenitors and mature lymphoid and myeloid cells. In fetal MSC co-cultures, PreProB-progenitors (CD10-ve) acquire surface CD10 *in vitro*, giving rise to ProB-progenitors and B-cells (Figure 4a, supplemental Figure 3b). By contrast, almost all ProB-progenitors lose surface CD34 expression within 3 days of culture to become B-cells and no PreProB-progenitors are generated (Figure 4a, supplemental Figure 3b). Fetal BM ELP from the same samples give rise to myeloid, NK- and B-lineage cells (including PreProB- and ProB-progenitors and B-cells), on human fetal BM MSC (data not shown). These data support the model of B-progenitor differentiation that places PreProB-progenitors upstream of ProB-progenitors and downstream of ELP in the fetal BM B-cell differentiation hierarchy. Consistent with this, complete rearrangements of the IgH locus (VDJ) are not detected in LMPP, ELP and PreProB-progenitors, while incomplete IgH rearrangements (DJ) are detected in some ELP and PreProB-progenitors and DJ or VDJ rearrangements in all ProB-progenitors, as well as complete VDJ rearrangements in CD34-CD19+ B-cells (Figure 4b, supplemental Table 5). As previously described for fetal B-cells^{27,28}, fetal BM B-progenitors also demonstrated preferential usage of IGHD7-27 (supplemental Table 5).

Gene expression profiling confirms that PreProB-progenitors lie upstream of ProB-progenitors

To confirm the B-progenitor hierarchy at the molecular level and investigate the transcriptional changes during B-progenitor differentiation we first performed transcriptional profiling of index-sorted single fetal BM cells (Lin2-CD34+ HSPC and Lin2-CD34-CD19+ B-cells; gating strategy shown in supplemental Figure 4), by qRT-PCR using a custom 96-gene panel (supplemental Table 2). A diffusion map demonstrated a lymphoid differentiation trajectory from HSC/MPP to B-cells and a myeloid trajectory from HSC/MPP to myeloid/erythroid progenitors (Figure 4c, supplemental Figure 5a, b).

To further define the hierarchical relationship between fetal lymphoid progenitors, we index-sorted single HSC, MPP and lymphoid progenitors and B-cells from fetal BM using the same gating strategy used for functional assays (supplemental Figure 1), and performed single cell qRT-PCR using a customized lymphoid gene panel (supplemental Table 3). A diffusion map confirmed a lymphoid differentiation trajectory from HSC to B-cells with PreProB-progenitors lying upstream of ProB-progenitors and downstream of ELP (Figure 4d, supplemental Figure 6 a,b) supporting our functional data. Early B-lineage gene expression (*CD79A*, *VPREB1*, *EBF1*, *IL7RA*, *IGHM*) is first detected in LMPP and increases progressively in ELP and PreProB-progenitors consistent with earlier reports²⁸, while expression of B-cell-specific genes (*DNTT*, *PAX5*, *CD19*, *MS4A1/CD20*) is not detected until ELP (Figure 4e, Supplemental Figures 5b, 6b). Most B-cell genes (e.g. *VPREB1*, *MME/CD10*, *RAG1*, *RAG2*, *CD72*, *CD9*, *LEF1* and *TCF3*) are expressed at higher levels in ProB- compared to PreProB-progenitors, consistent with a PreProB- to ProB-progenitor hierarchy suggesting their importance in B-progenitor maturation.

Next, to identify global transcriptomic differences between PreProB- and ProB-progenitors and their upstream progenitors, we performed RNA-sequencing of flow-sorted fetal BM HSPC (100 cells/population). Principal component analysis (PCA) based on the top 500 variably expressed genes clearly separates PreProB- from ProB-progenitors (Figure 5a). PCA also showed that PreProB-progenitors cluster adjacent to ELP (Figure 5a) although differential expression (DE) analysis shows higher expression of myeloid (*CD15*), and T-cell (*CD7*, *CD3E*, *CD44*) genes in ELP (Figure 5b,c, supplemental Table 6), reflecting the differences in their functional output. DE analysis found 808 variably expressed genes (FDR<0.1; supplemental Table 7) between PreProB- and ProB-progenitors, consistent with the distinct identity of these two B-progenitor types at a transcriptional level (Figure 5b). B-cell genes were upregulated from PreProB- to ProB-progenitors supporting the single cell qRT-PCR data (Figure 5d, supplemental Figure 7a). *DNTT* and *RAG1* genes were expressed at highest level in B-progenitors before being downregulated in B-cells (supplemental Figure 7c). Interestingly, PreProB-progenitors continued to express stem cell- (*HOXA5*, *KIT*, *LIN28B*, *RUNX2*), myeloid- (*MPO*, *CSF1R*) and T-cell- (*CD244*, *CD7*, *CD3D*) associated genes (Figure 5d, supplemental Figure 7b) despite their lack of myeloid/T-cell output *in vitro* or *in vivo*. Similar to adult murine CLP²⁹ and human cord blood B-progenitors³⁰, this most likely reflects the concurrent expression of a strong B-lineage

differentiation program together with markedly lower level (8-20-fold) of expression of myeloid genes in fetal PreProB-progenitors compared to multilineage (LMPP/MPP) and myeloid (GMP) progenitors (supplemental Figure 7c, d), that by contrast readily give rise to myeloid cells *in vitro*.

Together, these data confirm the single cell qRT-PCR data placing PreProB-progenitors upstream of ProB-progenitors and downstream of ELP.

PreProB- and ProB-progenitors show subtle differences in chromatin accessibility

To investigate whether the transcriptomic differences between PreProB- and ProB-progenitors are linked to differences in chromatin accessibility, we next performed ATAC-Sequencing²⁵ on flow-sorted fetal BM PreProB- and ProB-progenitors (2,000 cells each from 3 biological samples). PreProB- and ProB-progenitors both show an open chromatin accessibility pattern either upstream of, or overlapping with, gene bodies of key B-cell development genes including *PAX5*, *CD79A*, *RAG1*, *MME/CD10*, *DNTT*, *TCL1A* and *IGH* (supplemental Figure 8). Interestingly, stem cell and myeloid genes expressed in PreProB-progenitors (*LIN28B*, *CSF1R*, *MPO*) (supplemental Figure 7b) demonstrate qualitatively higher chromatin accessibility in PreProB-progenitors compared to ProB-progenitors (supplemental Figure 8). Genome-wide threshold-based differential chromatin accessibility analysis revealed differences in chromatin accessibility between PreProB- and ProB-progenitors (FDR <0.05, Figure 5e). We detected 351 differentially accessible regions in close vicinity of 341 genes (supplemental Table 8). A threshold-free comparative approach comparing chromatin accessibility at promoter regions of the 808 DE genes identified by RNA-Sequencing (supplemental Table 9) showed higher chromatin accessibility associated with genes that are upregulated during PreProB to ProB transition (p=0.02, Figure 5f). Taken together, these data show that chromatin accessibility in fetal BM progenitors reflects the transcriptomic profile of these cells and is cell context-dependent, demonstrating subtle differences even between closely related B-progenitor populations.

PreProB-progenitors in adult BM are functionally similar to their counterparts in fetal BM but exhibit ontogeny-related transcriptomic differences

To address the specific relationship of PreProB-progenitors to fetal B-cell development, we next investigated the frequency and characteristics of these progenitors in adult BM (n=6). PreProB-progenitors, as well as ELP, were rare in adult BM ($0.53 \pm 0.24\%$ and $0.25 \pm 0.11\%$ of Lin2-CD34+ cells respectively) (Figure 2a, supplemental Figure 9a). We then asked if

flow sorted (supplemental Figure 9b) adult BM PreProB-progenitors were functionally and molecularly similar to or distinct from their fetal BM counterparts. Like fetal BM PreProB-progenitors, adult BM PreProB- (and ProB-) progenitors differentiate to B-cells *in vitro*, with no myeloid or NK output (supplemental Figure 9c) although adult BM PreProB-progenitors generated lower numbers of B-cells than fetal progenitors (Supplemental Figure 9d). Finally, we performed single cell qRT-PCR on flow-sorted adult BM HSPC (1150 cells from 2 biological samples) to directly compare the B-cell developmental hierarchy between adult and fetal BM and identify any transcriptional differences. Although diffusion maps suggest that adult BM PreProB-progenitors may be more closely related to ProB-progenitors (Supplemental Figure 10) than in fetal BM (Figure 4d), adult and fetal PreProB-progenitors express very similar levels of most B-lineage genes (*CD79A*, *MME*, *MS4A1*, *PAX5*; supplemental Figure 10b). However, interestingly, compared to adult PreProB-progenitors, fetal PreProB-progenitors express higher levels of genes involved in DNA recombination (*DNTT*, *RAG1*), some myeloid (*MPO*) and HSC/leukemia genes (*TCF3*, *LEF1*, *MEF2C*, *IL7RA*) supporting previous work showing differences in gene expression between fetal and adult BM cells²⁸. As well as known fetal genes (*LIN28B*), fetal PreProB-progenitors uniquely express several genes implicated in iALL, that are either not expressed at all (*KLRK1*, *PPP1R14A*), or expressed at relatively lower levels in adult counterparts (*BAZ2B*, *BCAT1*)³¹⁻³⁴ (supplemental Figure 10b).

DISCUSSION

As knowledge about fetal B-cell development largely derives from studies in mice^{35,36}, we set out to identify the molecular and functional characteristics of human B-progenitor development in fetal BM. Here we demonstrate, using paired samples, that although B-lymphopoiesis takes place concurrently in liver and BM, it is in fetal BM that B-progenitors undergo a rapid and dramatic expansion during the early second trimester so that they form almost half of total BM HSPC by 20 pcw. The high frequency of B-progenitors in fetal BM so early in life, and the rate at which their frequency increases, have not previously been recognized. We also show for the first time that the earliest B-lymphoid restricted fetal progenitors are CD10-ve PreProB-progenitors which lie upstream of a second committed CD10+ve ProB-progenitor, from which they are transcriptionally distinct.

Importantly, CD10- PreProB-progenitors are not reported in adult BM suggesting they might play a specifically important role in fetal development. This led us to compare adult and fetal BM samples using the same flow cytometric panel which revealed the presence of a very small population of PreProB-progenitors (median 0.47 % of Lin2-CD34+ cells). Like their fetal counterparts, adult BM PreProB-progenitors give rise only to B-lineage cells *in vitro* and express the expected B-lineage gene expression program. However, fetal PreProB-progenitors, display a distinct, ontogeny-related gene expression pattern which is not seen in adult PreProB-progenitors. Fetal PreProB-progenitors uniquely express known fetal genes, such as *LIN28B*, and several genes implicated in infant ALL (iALL) which have not previously been recognized as human fetal-specific hematopoietic genes, such as *KLRK1* and *PPP1R14A*.

This raises the possibility that the identification of a CD10-ve B-progenitor in human FL and fetal BM brings us one step closer to unravelling the origins of B-progenitor leukemia in children, especially iALL. Although the target cell that undergoes transformation in iALL remains unclear³⁷⁻⁴¹, recent work suggests that for MLL-AF4+ iALL, it is likely to be a progenitor already committed to the B-lineage⁴² or a lymphoid-primed multipotent progenitor (LMPP)⁴¹. Fetal PreProB-progenitors display a number of characteristics that could provide a permissive molecular context for leukemic transformation by MLL rearrangements (MLLr). They are highly proliferative fetal-specific progenitors that have similar VDJ rearrangement status and immunophenotypic characteristics to iALL blasts⁴³, and display a similar transcriptome by RNA-sequencing (supplemental Figure 11a). They demonstrate accessible chromatin at the *IgH* gene locus, and expression of *DNTT* and *RAG1* genes, both of which might be important for fetal specific pre-BCR development and leukemia initiation^{7,44,45}. Furthermore, fetal PrePro-B progenitors, like iALL blasts, continue to express myeloid and stem cell gene programs (Figure 5d, supplemental Figure 7b), as well as MLLr iALL associated genes (supplemental Figure 11b). This is particularly pertinent given that iALL may undergo a lineage switch at relapse⁴⁶⁻⁴⁸. Our results raise the possibility of this occurring in a committed B-progenitor rather than an upstream multipotent progenitor, although further work will be required to support this.

Recent work identified another developmentally-restricted progenitor with B-lineage output in early FL in humans^{8,9}. This IL7R⁺ progenitor (defined as Lin-CD34⁺CD127⁺CD19⁻) has been suggested as a target cell for transformation by the leukemic fusion gene *ETV6-RUNX1*, one of the commonest causes of childhood CD10⁺ B-ALL. This has many similarities with the ELP we observed in fetal BM both at the functional level (B- and NK-cell as well as some myeloid output) and at the transcriptomic level. However, we also noted T-cell potential from fetal BM ELP, and show that these cells lie upstream of PreProB-progenitors consistent with their less restricted output. It is also possible that the ELP population is heterogeneous and contains within it a small subset of B-lineage restricted cells, however this could only be addressed by single cell in vitro assays which are technically extremely difficult given the rarity of this population. Interestingly, by PCA fetal BM ELP, like PreProB-progenitors, cluster closely with iALL blasts (supplemental Figure 11a) and might also serve as a permissive cell for MLLr transformation.

In summary, we describe the molecular and functional characteristics of B-progenitor development in human fetal BM for the first time. This reveals two types of B-lymphoid restricted progenitors, 'fetal-type' CD10^{-ve} PreProB-progenitors, which lie upstream of CD10⁺ ProB-progenitors from which they are molecularly distinct. PreProB-progenitors are very rare in adult BM, and differ from their fetal counterparts functionally and molecularly. Fetal PreProB-progenitors have a unique fetal-specific gene expression program that is absent in postnatal life and providing potential clues to the *in utero* origins of infant/childhood leukemia. This, along with their immunophenotypic and transcriptomic similarities with CD10^{-ve} iALL blasts and rapid expansion in fetal BM within a narrow developmental time window may indicate a permissive cell-intrinsic context and microenvironment for leukemia initiation although further studies are needed to investigate this directly.

Acknowledgements

A.R. is supported by a Bloodwise Clinician Scientist Fellowship (grants: 14041 and 17001), Lady Tata Memorial International Fellowship, and EHA-ASH Translational Research Training in Hematology Fellowship. I.R. is supported by the NIHR Oxford BRC and by a Bloodwise Program Grant (13001). T.A.M is supported by a Medical Research Council (UK) Molecular Haematology Unit grant MC_UU_12009/6 and by

the NIHR Oxford BRC. A.M. is supported by the NIHR Oxford BRC and MRC Senior Clinical Fellowship (MR/L006340/1). P.V. is supported by the MRC (MHU Award G1000729 and MRC Disease Team Award 4050189188, Single Cell Award MR/M00919X/1), CRUK (Program Grant C7893/A12796, Development Fund CRUKDF0176-DK and; CRUK Development Fund C5255/A20758. D.J.H.F.K. is supported by a CIHR Postdoctoral Fellowship. We thank the Gold September Charity for funding the ATAC-sequencing and The John Fell Fund for funding equipment. We would like to acknowledge the WIMM Flow Cytometry Facility which is supported by the MRC HIU; MRC MHU (MC_UU_12009); NIHR Oxford BRC; Kay Kendall Leukemia Fund (KKL1057), John Fell Fund (131/030 and 101/517), the EPA fund (CF182 and CF170) and by the WIMM Strategic Alliance awards G0902418 and MC_UU_12025. We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust Grant Reference 090532/Z/ 09/Z); the MRC WIMM Centre for Computational Biology (CCB), Radcliffe Department of Medicine, University of Oxford and ; the BMS, University of Oxford for their technical assistance. The human fetal material was provided by the Joint MRC/Wellcome Trust Grant 099175/Z/ 12/Z Human Developmental Biology Resource (<http://hdbr.org>). We thank Supat Thongjuea and Rajiv Mundayat for helpful advice on bioinformatic and statistical analyses and Dr Jack Bartram for advice on IgH rearrangement analysis.

Author Contributions

S.O.B., N.E., S.R., G.B., N.F., C.G., L.G., G.W., S.I., and A.R. performed the experiments and analyzed the data, N.C., D.K., B.Povinelli, A.A-D. C.B., I.V. and H.K. analyzed the data, B.Psaila., P.H., S.W., and P.B. provided samples, A.K., A.M., P.A., P.M., P.V. reviewed and edited the paper, S.O.B., T.A.M., I.R. and A.R. wrote and edited the paper, T.A.M., I.R. and A.R. conceptualized the project.

Disclosure of Conflicts of Interest

T.A.M. is one of the founding shareholders of Oxstem Oncology, a subsidiary company of OxStem Ltd (2016). All other authors declare no competing financial interests.

REFERENCES

1. Asma GE, Langlois van den Bergh R, Vossen JM. Development of pre-B and B lymphocytes in the human fetus. *Clin Exp Immunol*. 1984;56(2):407-414.
2. Bueno C, van Roon EH, Munoz-Lopez A, et al. Immunophenotypic analysis and quantification of B-1 and B-2 B cells during human fetal hematopoietic development. *Leukemia*. 2016;30(7):1603-1606.
3. Uckun FM, Ledbetter JA. Immunobiologic differences between normal and leukemic human B-cell precursors. *Proc Natl Acad Sci U S A*. 1988;85(22):8603-8607.
4. Nunez C, Nishimoto N, Gartland GL, et al. B cells are generated throughout life in humans. *J Immunol*. 1996;156(2):866-872.
5. Charbord P, Tavian M, Humeau L, Peault B. Early ontogeny of the human marrow from long bones: an immunohistochemical study of hematopoiesis and its microenvironment. *Blood*. 1996;87(10):4109-4119.
6. Grumayer ER, Griesinger F, Hummell DS, Brunning RD, Kersey JH. Identification of novel B-lineage cells in human fetal bone marrow that coexpress CD7. *Blood*. 1991;77(1):64-68.
7. Greaves M. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nat Rev Cancer*. 2018;18(8):471-484.
8. Alhaj Hussen K, Vu Manh TP, Guimiot F, et al. Molecular and Functional Characterization of Lymphoid Progenitor Subsets Reveals a Bipartite Architecture of Human Lymphopoiesis. *Immunity*. 2017;47(4):680-696 e688.
9. Boiers C, Richardson SE, Laycock E, et al. A Human IPS Model Implicates Embryonic B-Myeloid Fate Restriction as Developmental Susceptibility to B Acute Lymphoblastic Leukemia-Associated ETV6-RUNX1. *Dev Cell*. 2018;44(3):362-377 e367.
10. Haddad R, Guimiot F, Six E, et al. Dynamics of thymus-colonizing cells during human development. *Immunity*. 2006;24(2):217-230.
11. Blom B, Spits H. Development of human lymphoid cells. *Annu Rev Immunol*. 2006;24:287-320.
12. Sanz E, Munoz AN, Monserrat J, et al. Ordering human CD34+CD10-CD19+ pre/pro-B-cell and CD19- common lymphoid progenitor stages in two pro-B-cell development pathways. *Proc Natl Acad Sci U S A*. 2010;107(13):5925-5930.

- 510 13. Sanz E, Alvarez-Mon M, Martinez AC, de la Hera A. Human cord blood
511 CD34+Pax-5+ B-cell progenitors: single-cell analyses of their gene expression
512 profiles. *Blood*. 2003;101(9):3424-3430.
- 513 14. Roy A, Cowan G, Mead AJ, et al. Perturbation of fetal liver hematopoietic
514 stem and progenitor cell development by trisomy 21. *Proc Natl Acad Sci U S A*.
515 2012;109(43):17579-17584.
- 516 15. Agraz-Doblas A, Bueno C, Bashford-Rogers R, et al. Unravelling the cellular
517 origin and clinical prognostic markers of infant B-cell acute lymphoblastic leukemia
518 using genome-wide analysis. *Haematologica*. 2019.
- 519 16. Andersson AK, Ma J, Wang J, et al. The landscape of somatic mutations in
520 infant MLL-rearranged acute lymphoblastic leukemias. *Nat Genet*. 2015;47(4):330-
521 337.
- 522 17. Awong G, La Motte-Mohs RN, Zuniga-Pflucker JC. In vitro human T cell
523 development directed by notch-ligand interactions. *Methods Mol Biol*. 2008;430:135-
524 142.
- 525 18. Dominici M, Le Blanc K, Mueller I, et al. Minimal criteria for defining
526 multipotent mesenchymal stromal cells. The International Society for Cellular
527 Therapy position statement. *Cytotherapy*. 2006;8(4):315-317.
- 528 19. Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S, Sandberg R.
529 Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*. 2014;9(1):171-
530 181.
- 531 20. Psaila B, Barkas N, Iskander D, et al. Single-cell profiling of human
532 megakaryocyte-erythroid progenitors identifies distinct megakaryocyte and erythroid
533 differentiation pathways. *Genome Biol*. 2016;17:83.
- 534 21. Angerer P, Haghverdi L, Buttner M, Theis FJ, Marr C, Buettner F. destiny:
535 diffusion maps for large-scale single-cell data in R. *Bioinformatics*. 2016;32(8):1241-
536 1243.
- 537 22. Haghverdi L, Buttner M, Wolf FA, Buettner F, Theis FJ. Diffusion
538 pseudotime robustly reconstructs lineage branching. *Nat Methods*. 2016;13(10):845-
539 848.
- 540 23. van Dongen JJ, Langerak AW, Bruggemann M, et al. Design and
541 standardization of PCR primers and protocols for detection of clonal immunoglobulin
542 and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the
543 BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 2003;17(12):2257-
544 2317.

- 545 24. Brochet X, Lefranc MP, Giudicelli V. IMGT/V-QUEST: the highly
546 customized and integrated system for IG and TR standardized V-J and V-D-J
547 sequence analysis. *Nucleic Acids Res.* 2008;36(Web Server issue):W503-508.
- 548 25. Corces MR, Buenrostro JD, Wu B, et al. Lineage-specific and single-cell
549 chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat*
550 *Genet.* 2016;48(10):1193-1203.
- 551 26. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for
552 Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol.* 2015;109:21
553 29 21-29.
- 554 27. Roy A, Bystry V, Bohn G, et al. High resolution IgH repertoire analysis
555 reveals fetal liver as the likely origin of life-long, innate B lymphopoiesis in humans.
556 *Clin Immunol.* 2017;183:8-16.
- 557 28. Bertrand FE, 3rd, Billips LG, Burrows PD, Gartland GL, Kubagawa H,
558 Schroeder HW, Jr. Ig D(H) gene segment transcription and rearrangement before
559 surface expression of the pan-B-cell marker CD19 in normal human bone marrow.
560 *Blood.* 1997;90(2):736-744.
- 561 29. Hsu CL, King-Fleischman AG, Lai AY, Matsumoto Y, Weissman IL, Kondo
562 M. Antagonistic effect of CCAAT enhancer-binding protein- α and Pax5 in
563 myeloid or lymphoid lineage choice in common lymphoid progenitors. *Proc Natl*
564 *Acad Sci U S A.* 2006;103(3):672-677.
- 565 30. Laurenti E, Doulatov S, Zandi S, et al. The transcriptional architecture of early
566 human hematopoiesis identifies multilevel control of lymphoid commitment. *Nat*
567 *Immunol.* 2013;14(7):756-763.
- 568 31. Armstrong SA, Staunton JE, Silverman LB, et al. MLL translocations specify
569 a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet.*
570 2002;30(1):41-47.
- 571 32. Trentin L, Giordan M, Dingermann T, Basso G, Te Kronnie G, Marschalek R.
572 Two independent gene signatures in pediatric t(4;11) acute lymphoblastic leukemia
573 patients. *Eur J Haematol.* 2009;83(5):406-419.
- 574 33. Kerry J, Godfrey L, Repapi E, et al. MLL-AF4 Spreading Identifies Binding
575 Sites that Are Distinct from Super-Enhancers and that Govern Sensitivity to DOT1L
576 Inhibition in Leukemia. *Cell Rep.* 2017;18(2):482-495.
- 577 34. Torelli GF, Peragine N, Raponi S, et al. Recognition of adult and pediatric
578 acute lymphoblastic leukemia blasts by natural killer cells. *Haematologica.*
579 2014;99(7):1248-1254.

580 35. Montecino-Rodriguez E, Fice M, Casero D, Berent-Maoz B, Barber CL,
581 Dorshkind K. Distinct Genetic Networks Orchestrate the Emergence of Specific
582 Waves of Fetal and Adult B-1 and B-2 Development. *Immunity*. 2016;45(3):527-539.

583 36. Li YS, Zhou Y, Tang L, Shinton SA, Hayakawa K, Hardy RR. A
584 developmental switch between fetal and adult B lymphopoiesis. *Ann N Y Acad Sci*.
585 2015;1362:8-15.

586 37. Bardini M, Woll PS, Corral L, et al. Clonal variegation and dynamic
587 competition of leukemia-initiating cells in infant acute lymphoblastic leukemia with
588 MLL rearrangement. *Leukemia*. 2015;29(1):38-50.

589 38. Castor A, Nilsson L, Astrand-Grundstrom I, et al. Distinct patterns of
590 hematopoietic stem cell involvement in acute lymphoblastic leukemia. *Nat Med*.
591 2005;11(6):630-637.

592 39. Hotfilder M, Rottgers S, Rosemann A, et al. Leukemic stem cells in childhood
593 high-risk ALL/t(9;22) and t(4;11) are present in primitive lymphoid-restricted
594 CD34+CD19- cells. *Cancer Res*. 2005;65(4):1442-1449.

595 40. le Viseur C, Hotfilder M, Bomken S, et al. In childhood acute lymphoblastic
596 leukemia, blasts at different stages of immunophenotypic maturation have stem cell
597 properties. *Cancer Cell*. 2008;14(1):47-58.

598 41. Malouf C, Ottersbach K. The fetal liver lymphoid-primed multipotent
599 progenitor provides the prerequisites for the initiation of t(4;11) MLL-AF4 infant
600 leukemia. *Haematologica*. 2018.

601 42. Lin S, Luo RT, Shrestha M, Thirman MJ, Mulloy JC. The full transforming
602 capacity of MLL-Af4 is interlinked with lymphoid lineage commitment. *Blood*.
603 2017;130(7):903-907.

604 43. Jansen MW, Corral L, van der Velden VH, et al. Immunobiological diversity
605 in infant acute lymphoblastic leukemia is related to the occurrence and type of MLL
606 gene rearrangement. *Leukemia*. 2007;21(4):633-641.

607 44. Kirkham CM, Scott JNF, Wang X, et al. Cut-and-Run: A Distinct Mechanism
608 by which V(D)J Recombination Causes Genome Instability. *Mol Cell*.
609 2019;74(3):584-597 e589.

610 45. Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the
611 predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute
612 lymphoblastic leukemia. *Nat Genet*. 2014;46(2):116-125.

613 46. Jiang JG, Roman E, Nandula SV, Murty VV, Bhagat G, Alobeid B.
614 Congenital MLL-positive B-cell acute lymphoblastic leukemia (B-ALL) switched
615 lineage at relapse to acute myelocytic leukemia (AML) with persistent t(4;11) and

616 t(1;6) translocations and JH gene rearrangement. *Leuk Lymphoma*. 2005;46(8):1223-
617 1227.

618 47. Rossi JG, Bernasconi AR, Alonso CN, et al. Lineage switch in childhood
619 acute leukemia: an unusual event with poor outcome. *Am J Hematol*. 2012;87(9):890-
620 897.

621 48. Sakaki H, Kanegane H, Nomura K, et al. Early lineage switch in an infant
622 acute lymphoblastic leukemia. *Int J Hematol*. 2009;90(5):653-655.

623 49. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq
624 aligner. *Bioinformatics*. 2013;29(1):15-21.

625 50. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose
626 program for assigning sequence reads to genomic features. *Bioinformatics*.
627 2014;30(7):923-930.

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

630 52. Hay D, Hughes JR, Babbs C, et al. Genetic dissection of the alpha-globin
631 super-enhancer in vivo. *Nat Genet*. 2016;48(8):895-903.

632 53. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC.
633 *Genome Res*. 2002;12(6):996-1006.

634 54. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment
635 using MACS. *Nat Protoc*. 2012;7(9):1728-1740.

636 55. Ross-Innes CS, Stark R, Teschendorff AE, et al. Differential oestrogen
637 receptor binding is associated with clinical outcome in breast cancer. *Nature*.
638 2012;481(7381):389-393.

639 56. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable
640 read mapping by seed-and-vote. *Nucleic Acids Res*. 2013;41(10):e108.
641

FIGURE LEGENDS:

Figure 1: High frequency of B progenitors in first and second trimester fetal bone marrow (FBM)

a) Frequency of fetal liver (FL) and fetal BM HSC/MPP and b) LMPP from 6-21 **post-conception weeks (pcw)**. Data expressed as % of Lin-CD34+ (mean \pm SEM, n=59 for FL and n=32-35 for fetal BM).

c) Frequency of FL and fetal BM early lymphoid progenitors (ELP), d) PreProB-progenitors and e) ProB-progenitors from 6-21 pcw. Data expressed as % of Lin2-CD34+ (mean \pm SEM, n=39 for FL and n=32-38 for fetal BM).

f) Frequency of FL and fetal BM B-cells (CD34-CD19+) from 6-21 pcw. Data expressed as % of CD34neg (n=37 for FL and n=32-38 for fetal BM).

g) Representative flow-cytometric analysis of fetal BM Lin2-CD34+ cells showing the gating strategy used to identify CD10-CD34+CD19+ PreProB-progenitors and CD10+CD34+CD19+ ProB-progenitors.

h) Frequency of fetal BM hematopoietic stem and multipotent progenitors (HSC/MPP), lymphoid primed multipotent progenitors (LMPP), ELP, PreProB-progenitors and ProB-progenitors from 8 to 21 pcw. Data expressed as % of Lin2-CD34+ cells for B-progenitors and Lin-CD34+ cells for ELP, HSC/MPP/LMPP (mean \pm SEM; n= 32-38).

i) Frequency of second trimester (12-21 weeks) FL and fetal BM hematopoietic stem and progenitor cell (HSPC) populations expressed as % of Lin-/Lin2-CD34+ cells (mean \pm SEM, n=17-49 for FL and n=16-27 for fetal BM).

j) Matched FL and fetal BM from the same biological samples, showing mean frequency of PreProB- and ProB-progenitors expressed as % of Lin2-CD34+ cells. Individual samples represented as circles for FL and squares for fetal BM.

Lin: CD2, CD3, CD14, CD16, CD56, CD19, CD235a

Lin2: CD2, CD3, CD14, CD16, CD56, CD235a

All graphs show mean \pm SEM where n \geq 3. Statistical significance determined by two-way ANOVA. *p<0.05; ** p<0.01; ***p<0.001; ****p<0.0001.

Figure 2: PreProB- progenitors are enriched in fetal life and B-progenitor expansion is favored over differentiation in a narrow developmental time window.

- a) Frequency of ELP, PreProB- and ProB-progenitors in the Lin2-CD34+ compartment of second trimester fetal BM (n=29-32, 12-21pcw) and adult BM (n=6). Representative flow plots for PreProB- and ProB-progenitors are shown on the right. Statistical significance determined by unpaired t-test: *p<0.05; ***p<0.001; ****p<0.0001
- b) Cell cycle analysis by flow cytometry: representative flow plots of fetal BM HSPC populations showing the proportion of cells in G0, G1 and S-G2-M phases.
- c) Percentage of cells in S-G2-M phase in each fetal BM HSPC population. Data represented as mean + SEM (n=3). Statistical significance determined by paired t-test.
- d) Changes in frequency of fetal BM PreProB-progenitors, ProB-progenitors and B-cells from 8-21 weeks' pcw. B-cells are shown as % of CD34- cells and B-progenitors as % of Lin2-CD34+ (n=32-38). Shaded area highlights the developmental time window when B-progenitor expansion occurs prior to B-cell differentiation. Lin2 antibody cocktail: CD2,CD3,CD14,CD16,CD56,CD235a

Figure 3: Fetal BM PreProB- and ProB-progenitors are committed to the B-lymphoid lineage

- a) Cell frequency at day 14 and day 21 after co-culturing flow-sorted fetal BM HSPC on murine MS-5 stromal cells with cytokines (FLT3L, SCF, IL7 and IL2). Data shown as % of human CD45+ cells (mean - SEM, n=3-5).
- b) Frequency of T cells (CD4+ and/or CD8+) produced by day 7 and day 14 after co-culturing flow-sorted fetal BM HSPC on OP9-DL1 (GFP+) stroma with cytokines (FLT3L and IL7). Data shown as % GFP- human cells (mean + SEM, n=4).
- c) B-cell differentiation assay of flow-sorted ELP co-cultured on MS-5 stromal cells with cytokines (FLT3L, SCF, IL7 and IL2) showing output on day 3, 7 and 14 (results expressed as mean - SEM, n=2-4).
- d) Myeloid clonogenic assays showing day 14 colony readout from flow-sorted fetal BM HSPC in methylcellulose colony assays (data shown as mean + SEM, n=4).
- (e) Engraftment of fetal BM Lin2-CD34+CD19-CD10- progenitors, PreProB-progenitors and ProB-progenitors in the BM of NSG mice 2-3 weeks after transplantation (n=4 for each progenitor population, total n=12). Data shown as a % human CD45 (hCD45) cells of all CD45+ cells in mouse BM (mean ± SEM).

(f) Identity of hCD45⁺ cells engrafted in mouse BM 2-3 weeks post-transplant. The plot in the centre shows the breakdown of all human CD45⁺ cells and the plots either side of the central plot detail the progenitors identified in human CD45⁺CD34⁺ cells.

Figure 4: Fetal BM CD10⁻ PreProB-progenitors lie upstream of CD10⁺ ProB-progenitors in the fetal B-cell developmental hierarchy

- a) B-cell differentiation assay of flow-sorted PreProB- and ProB-progenitors co-cultured on fetal BM MSC with cytokines (FLT3L, SCF, IL7 and IL2); showing proportions of progenitors and mature B-cells produced by day 3, 7 and 14 (results expressed as mean -SEM, n=3).
- b) Pattern of IgH gene rearrangement detected in whole genome amplified DNA from flow-sorted fetal BM HSPC expressed as % of samples tested (n=3-7 for each subpopulation)
- (c) Diffusion map of single cell qRT-PCR (total of 938 cells from 2 biological samples) of index-sorted Lin2-CD34⁺ or Lin2-CD34-CD19⁺ fetal BM cells showing B-lymphoid and myeloid/erythroid differentiation trajectories. Genes and TaqMan probes detailed in supplemental Table 2.
- (d) Diffusion map of single cell qRT-PCR (total of 1400 cells from 3 biological samples) of flow-sorted fetal BM HSPC and B-cells showing differentiation trajectory from HSC through to B-cells. Genes and TaqMan probes detailed in supplemental Table 3.
- (e) Normalized gene expression of selected B-cell associated genes in flow-sorted fetal BM HSPC.

Figure 5: Fetal BM PreProB-progenitors are molecularly distinct from ELP and ProB-progenitors

- a) Bulk RNA-sequencing of 100 cells from each population (n=3): principal component analysis (PCA) of each HSPC population using the top 500 most variable genes
- b) Heatmap summarizing differentially expressed (DE) genes (FDR<0.1) between fetal BM ELP and PreProB-progenitors (292 DE genes) and PreProB- and ProB-progenitors (808 DE genes). Genes detailed in supplemental Tables 6 and 7.
- c) Log2 fold change of selected lymphoid and myeloid genes that are significantly differentially expressed between fetal BM ELP and PreProB-progenitors.

- d) Log2 fold change of selected lymphoid, myeloid and stem cell genes that are significantly differentially expressed between fetal BM PreProB- and ProB-progenitors.
- e) Differential chromatin accessibility analysis of ATAC-sequencing from PreProB- and ProB-progenitors summarized by PCA and heatmap. Details of differentially accessible genes in supplemental Table 8.
- f) Using ATAC-Sequencing, chromatin was more accessible around the promoter regions of genes differentially expressed between PreProB- and ProB-progenitors ($p=0.027$, $n=2$).