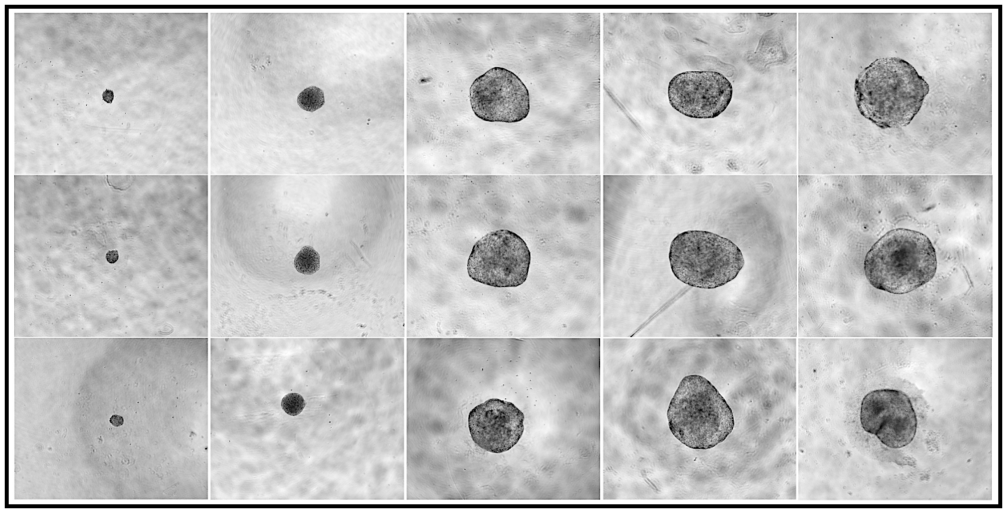

# The Role of Wnt Signaling in an Embryoid Body Generation Protocol for Hematopoiesis Studies

## Investigating the Emergence of Endothelial and Hematopoietic Stem and Progenitor Cells in hiPSC-Derived Embryoid Bodies Through RT-qPCR and Flow Cytometry Analysis

Master's Thesis in Biotechnology

Conducted at Sahlgrenska Center for Cancer Research

Group Guibentif, University of Gothenburg

Elin Hilpold Berntsson

DEPARTMENT OF MATHEMATICAL SCIENCES

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024
www.chalmers.se

Master's thesis 2024

# The Role of Wnt Signaling in an Embryoid Body Generation Protocol for Hematopoiesis Studies

Investigating the Emergence of Endothelial and Hematopoietic Stem and Progenitor Cells in hiPSC-Derived Embryoid Bodies Through RT-qPCR and Flow Cytometry Analysis

Elin Hilpold Berntsson


CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Mathematical Sciences
Chalmers University of Technology
*Conducted at Sahlgrenska Center for Cancer Research*
Group Guibentif, University of Gothenburg
Gothenburg, Sweden 2024

The Role of Wnt Signaling in an Embryoid Body Generation Protocol for Hematopoiesis Studies
*Investigating the Emergence of Endothelial and Hematopoietic Stem and Progenitor Cells in hiPSC-Derived Embryoid Bodies Through RT-qPCR and Flow Cytometry Analysis*
ELIN HILPOLD BERNTSSON



Supervisor: Carolina Guibentif, Sahlgrenska Center for Cancer Research
Examiner: Erik Kristiansson, Department of Mathematical Statistics

Master's Thesis 2024
Department of Mathematical Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Sahlgrenska Center for Cancer Research
Group Guibentif, University of Gothenburg
SE-413 90 Gothenburg

Cover: Microscopy images of generated hiPSC-derived embryoid bodies

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2024

The Role of Wnt Signaling in an Embryoid Body Generation Protocol for Hematopoiesis Studies
*Investigating the Emergence of Endothelial and Hematopoietic Stem and Progenitor Cells in hiPSC-Derived Embryoid Bodies Through RT-qPCR and Flow Cytometry Analysis*
ELIN HILPOLD BERNTSSON
Department of Mathematical Sciences
Chalmers University of Technology

# Abstract

Research indicates that childhood leukemia originates *in utero* during hematopoietic development, emphasizing the importance of understanding embryonic hematopoiesis to identify disruptions leading to diseases such as leukemia. However, studying human embryonic hematopoiesis *in vivo* faces significant ethical and technical challenges. As an alternative, embryoid bodies (EBs) derived from human induced pluripotent stem cells (hiPSCs) offer a promising *in vitro* model for studying human developmental processes and hematological diseases. This thesis evaluates an established EB generation protocol for its efficacy in differentiating hiPSCs into hematopoietic stem and progenitor cells (HSPCs). Through the application of real-time quantitative PCR, this study investigates the activation of the Wnt signaling pathway, marked by *RSPO3* gene expression, assessing its impact on cell fate decisions within the EBs. Key findings reveal that activation of the Wnt pathway not only induces the emergence of endothelial cells but also facilitates the specification of HSPCs. The robust upregulation of genes such as *SOX17, CDH5, RUNX1, GATA2,* and *HOXA9* over time validates the progressive development of endothelial and hematopoietic lineages. Flow cytometry analysis results also confirm the emergence of endothelial and blood cell populations. However, the response of differentiating hiPSCs to Wnt signaling varies between independent experiments, suggesting that further optimization of the EB generation protocol can be implemented. Overall, the results underscore the role of Wnt signaling in lineage specification, offering valuable insights into hematopoietic differentiation and development in a controlled environment.

Keywords: hiPSC, hemogenic endothelium, HSPC, embryoid bodies, Wnt signaling, RSPO3

# Acknowledgments

I would like to express my gratitude to the following individuals, whose support has been invaluable to my academic journey.

Special thanks to everyone in Group Carolina Guibentif, especially my supervisor Dr. Carolina Guibentif, whose indispensable guidance and involvement have steered me in the right direction throughout this project. I am particularly grateful to Anuntxi Monsalve for patiently answering all my questions, providing substantial support, and guiding me in the cell culture lab. My appreciation also extends to Dr. Pavan Kumar for his insightful answers and assistance with the project's methods. Thanks to Dr. Anna Hogmalm for her readiness to help whenever needed, and Iman Ouadria, for providing the samples that enabled me to learn RT-qPCR. I am also thankful to Sofia Nilsson for her support at the project's initiation and Dr. Magdalena Strauss for her help with the statistical analysis.

My gratitude goes to Group Pekka Jaako, especially Daniel Sjövall, for his help with the analysis of the leukemia dataset.

A big thank you to Palmqvist Lab for sharing the ChiPSC22 cell line, which was crucial for the generation of embryoid bodies.

My appreciation goes to Professor Erik Kristiansson, whose assistance throughout this thesis has offered me calm and clarity.

Lastly, I owe my deepest thanks to my family and friends. Your unwavering support and love have been the cornerstone of not only my studies but also my life. I could not have reached this point without you.

Thank you all!

Elin Hilpold Berntsson, Gothenburg, May 2024

# List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

| | |
|---|---|
| A-P | Anterior-Posterior |
| AGM | Aorta-Gonad-Mesonephros |
| ALL | Acute Lymphocytic Leukemia |
| AML | Acute Myelogenous Leukemia |
| ANOVA | Analysis of Variance |
| bFGF | Basic Fibroblast Growth Factor |
| BPT | Before Pre-Treatment |
| CHIR | CHIR99021 |
| CLL | Chronic Lymphocytic Leukemia |
| CML | Chronic Myelogenous Leukemia |
| D | Day |
| DMSO | Dimethyl Sulfoxide |
| DOX | Doxycycline |
| DPBS | Dulbecco's Phosphate Buffered Saline |
| E6 | Essential 6 |
| E | Embryonic Day |
| EB | Embryoid Body |
| EHT | Endothelial-to-Hematopoietic Transition |
| EMP | Erythromyeloid Progenitor |
| EMT | Epithelial-to-Mesenchymal Transition |
| Exp | Experiment(s) |
| FMO | Fluorescence Minus One |
| GSK3 | Glycogen Synthase Kinase 3 |
| HSC | Hematopoietic Stem Cell |
| HSPC | Hematopoietic Stem and Progenitor Cell |
| hiPSC | Human Induced Pluripotent Stem Cell |
| iPSC | Induced Pluripotent Stem Cell |
| LRP | Lipoprotein Receptor-Related Protein |
| MLL | Histone Lysine Methyltransferase 2A |
| NS | NutriStem |
| PT | Pre-Treatment |
| RT-qPCR | Real-time Quantitative Polymerase Chain Reaction |
| SD | Standard Deviation |
| TrypLe | TrypLe Select Enzyme |
| UMAP | Uniform Manifold Approximation and Projection |
| VEGF | Vascular Endothelial Growth Factor |

# Contents

# List of Figures

# List of Tables

# 1

# Introduction

This thesis begins with a background that explores different aspects central to the project. The background sets the stage for a detailed problem statement that highlights current challenges in the field, particularly those related to the use of Wnt signaling in the differentiation of human induced pluripotent stem cells (hiPSCs). Following the problem statement, the aims of the project are outlined, detailing the specific objectives intended to address these challenges and advance our understanding of hematopoietic development.

## 1.1 Background

In this section, relevant background to establish the context of this research project is presented. An overview of the current understanding of the hematologic disease leukemia is given, focusing on its etiology and prevalence, specifically in children. This is followed by an outline of embryogenesis, emphasizing embryonic hematopoiesis which is a critical process for understanding hematologic development and disorders. I will then address various methodologies used to study hematopoiesis, specifically the use of hiPSCs, highlighting key factors that play important roles in this field of research.

### 1.1.1 Leukemia: Global Impact and Childhood Leukemia Initiation

Leukemia is a disease group that encompasses different variants of cancers of the white blood cells, known as leukocytes [1]. These variants are divided into different groups, depending on how rapidly the abnormal leukocytes proliferate and what origin they have. The more aggressively developing kind is called acute, and its emergence is thereby quite sudden and unexpected. The second type, chronic leukemia, emerges gradually while progressing slowly over months, or even years. Regarding the origin, there are lymphocytic leukemia and myelogenous leukemia which start in the lymphoid cells and myeloid cells, respectively. As such, the most common subtypes are chronic lymphocytic leukemia (CLL), acute myelogenous leukemia (AML), chronic myelogenous leukemia (CML), and acute lymphocytic leukemia (ALL), listed in descending order of prevalence. There are additional leukemia types, for example T-cell leukemia which arises from mature white blood cells, but these are not as common as the aforementioned variants.

In 2020, approximately 2.5% of all new cancer incidences were attributed to leukemia, which amounts to a total of 474,519 new leukemia cases [2]. As for the global mortality rate that year, 3.1% of cancer-related deaths were due to leukemia. It's worth noting that, when comparing the age-standardized rates, statistics show that the worldwide variation of incidence was almost five-fold. For instance, countries with higher income levels had a higher incidence [2]. However, this could in part be a result of under-reporting of leukemia cases in developing countries. While

leukemia affects a broad demographic, it's particularly concerning in younger populations. Each year, around 400,000 new cancer cases are diagnosed in children and teenagers (0-19 years) worldwide [3]. The most prevalent types of leukemia in children are the acute variants, while chronic leukemias are rarer. Given that leukemia is the predominant cancer type in children (0-14 years), and cancer is one of the leading causes of child mortality worldwide [4], [5], there is an urgency to understand childhood leukemia.

#### *In Utero* Initiation of Childhood Leukemia

Ever since the hypothesis that childhood leukemia originates prenatally was proposed, supporting evidence has accumulated [6]–[8]. Central to the investigation is ALL, the most prevalent form of childhood leukemia [8]. Studies supporting the hypothesis of an *in utero* emergence of ALL include research on identical twins, both of whom developed leukemia. Identical and unique clonal markers, specifically chromosomal translocations, were identified in the leukemic cells of both siblings. This finding indicates a prenatal initiation of the disease. A common translocation involves the histone lysine methyltransferase 2A (MLL) gene, associated with both ALL and AML [9]. This translocation has a high incident rate in infants, occurring in approximately 50% of AML cases and 70-80% of ALL cases [6]. Furthermore, studies analyzing for instance birth blood spots have provided additional insights into the prenatal origins of this disease by identifying leukemia-associated genetic changes [7], [8].

Although these findings suggest that the initial genetic events leading to acute leukemia occur *in utero*, the precise mechanisms and specific origins of these mutations during hematopoietic development remain elusive. Hence, this area of research continues to expand and evolve, underscoring the need for ongoing investigations.

### 1.1.2 Embryogenesis Encompasses Various Processes

Embryogenesis begins when a sperm and an oocyte fuse to create a diploid zygote [10], [11]. Following stages lead to the formation of the blastocyst by embryonic day (E) 5 in humans [11]. The blastocyst includes three specified lineages: the pluripotent epiblast, which is involved in forming the embryo proper, hence giving rise to all cell types in the body; the extraembryonic trophectoderm, which gives rise to the placenta; and the extra-embryonic hypoblast, which contributes to the yolk sac [10]–[12]. Both the pluripotent epiblast and the extra-embryonic hypoblast are derived from the apolar inner cell mass of the early blastocyst. Each of these three lineages plays a distinct role in subsequent embryonic and fetal development.

The knowledge about early embryogenesis pre-implantation is quite detailed, yet the understanding of human embryo development post-implantation is notably limited [13]. This is due to various reasons, such as ethical considerations and limitations in available models for studying embryogenesis. Despite this, historical analyses have provided insights into critical developmental stages. For instance, it has been observed that implanting blastocysts undergo key morphological changes, transforming into a disc-shaped embryo [14]. This is often referred to as the bilaminar embryonic disc and is made up of the pluripotent epiblast and the hypoblast [15]. This initiates the spatial organization of the dorsal/ventral axis, positioning the pluripotent epiblast dorsally relative to the hypoblast. Subsequent morphological changes involve the formation of the amniotic cavity and the primary yolk sac at approximately E8. At E10 the extra-embryonic mesoderm is formed derived from epiblast cells, and by migration leads to the enclosing of these structures.

Gastrulation is the process that directly succeeds the formation of the bilaminar embryonic disc and takes place in the third week of development [11], [13], [15]. Gastrulation gives rise to the definitive endoderm, intra-embryonic mesoderm, and ectoderm, establishing the trilaminar disc. The initiation of gastrulation is marked by the appearance of the primitive streak. This streak forms along the embryonic midline and acts as a reference for the subsequent migration of epiblasts, establishing the anterior-posterior (A-P) axis. These epiblast cells first undergo the epithelial-to-mesenchymal transition (EMT), to then migrate from the caudal (posterior) part of the embryo towards the anterior. As a result, the elongation of the embryo is initiated.

The axial elongation of the embryo is a process driven and stabilized by different signaling pathways, such as the Wnt pathway [16]. This pathway is involved in an extensive number of processes in embryogenesis [17]. Except for its role in gastrulation [18] by promoting EMT [19], [20] and the A-P axis definition [21], it is essential in the hematopoietic development, by for instance regulating hematopoietic stem and progenitor cells (HSPCs) [22].

The canonical Wnt/$\beta$-catenin signaling is the most studied branch of the Wnt pathway, where Wnt glycoproteins are secreted and bind to the Frizzled cell membrane receptor [21] (See Figure 1). This inhibits the degradation of the transcription co-activator $\beta$-catenin by a protein complex, including proteins such as glycogen synthase kinase 3 (GSK3) [23], thus accumulating $\beta$-catenin in the cytoplasm [21]. This enables the nuclear translocation of the transcription co-activator where its subsequent binding to the TCF/LEF protein complex initiates transcription of specific target genes, *C-MYC* for instance [24].

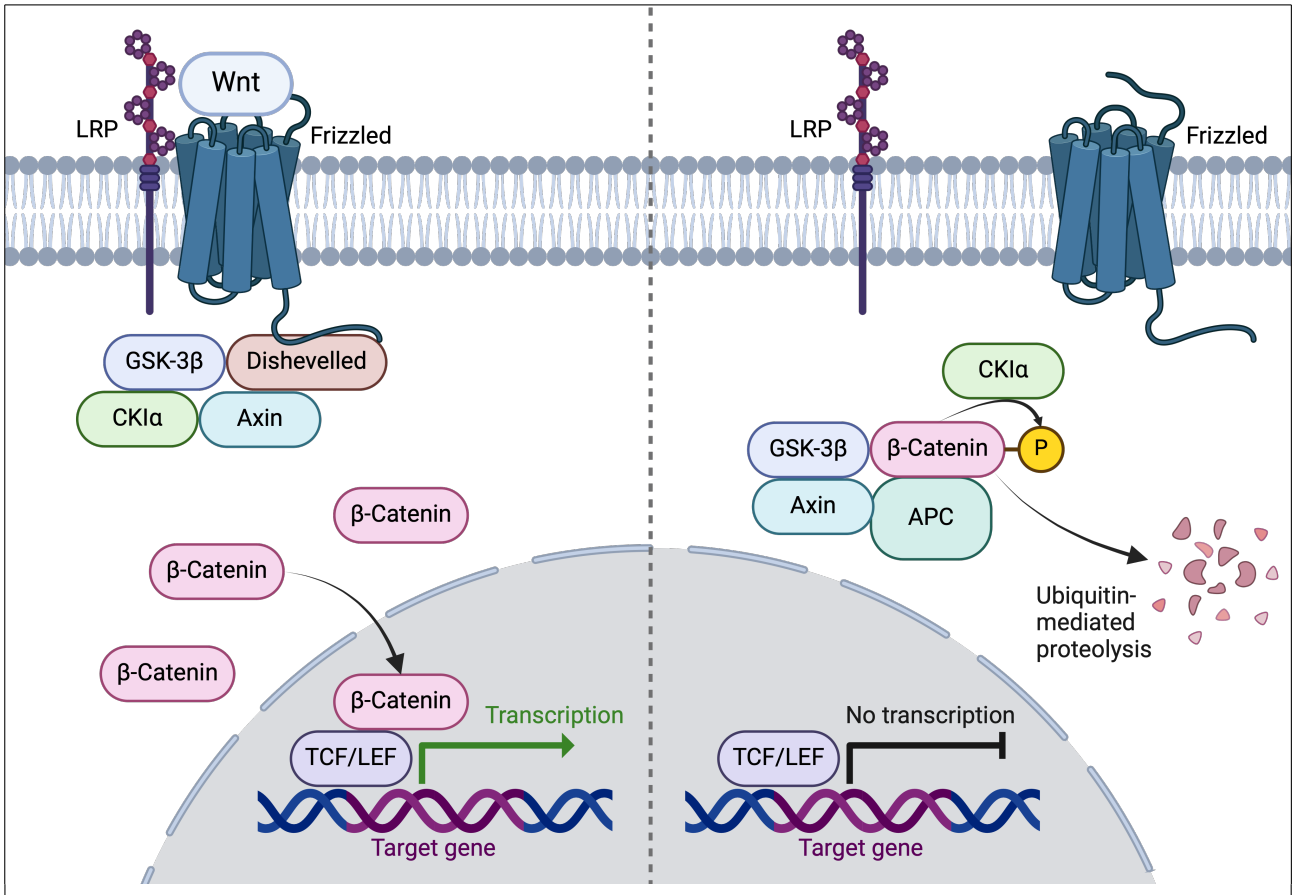


**Figure 1: The Canonical Wnt/$\beta$-Catenin Signaling Pathway**

Illustration of the Wnt pathway, either activated (left) or inactivated (right). When no Wnt glycoproteins bind to the Frizzled receptor, the pathway is inactivated and the protein complex, including the GSK3 protein, degrades $\beta$-catenin via ubiquitin-mediated proteolysis. When Wnt binds to the receptor, this degradation is inhibited.

Wnt signaling is enhanced by different proteins, such as RSPO3 which belongs to the R-Spondin protein family [25]. These proteins have also been shown to be coexpressed with and enhanced by Wnt proteins, suggesting a positive feedback modulation of the pathway [26].

In defining the A-P axis in humans and other vertebrates, the *HOX* genes have shown to be key factors [27]–[29]. This cluster of genes is involved in making sure structures form in the correct location and order, and it involves 38 genes, including the caudal transcription factor *HOXA9* [27], [28]. This gene has also been shown to play an important part in the early hematopoietic development and in involved HSPCs [30], [31].

In human organogenesis, the three germ layers derived from the gastrulation process give rise to all organs in the embryo proper [32]. Each germ layer is responsible for specific organs included in different organ systems. In particular, the mesoderm layer gives rise to cell types composing the musculoskeletal system, but it also contributes to the formation of the cardiovascular system. The mesoderm layer is thereby critical for the hematopoietic development in the embryo.

### Embryonic Hematopoiesis Occurs in Distinct Overlapping Waves

Understanding embryonic hematopoiesis is essential to determine the initiation of hematopoietic disruptions leading to diseases such as leukemia. The process of blood cell development initiates early in embryonic growth, occurring parallel to and interlinked with key developmental stages such as gastrulation, elongation, and organogenesis. Research has established that the hematopoietic system, in humans and other vertebrates, develops in distinct *primitive* and *definitive* waves [33], [34].

Primitive hematopoiesis initiates in the extra-embryonic yolk sac during gastrulation [33], [34]. This process predominantly generates primitive erythrocytes, macrophages, and megakaryocytes, derived from mesoderm progenitors [33], [35]. Primitive nucleated erythrocytes are crucial for distributing oxygen throughout the embryo, especially as it grows larger and distribution by simple diffusion is no longer sustainable. Macrophages and megakaryocytes are important for early tissue remodeling and the development of the immune system.

The pro-definitive wave is initiated by the emergence of erythromyeloid progenitors (EMPs) in the yolk sac's blood circulation [33], [36], [37]. The EMPs stem from a specialized subset of endothelial cells, hemogenic endothelium, in a process known as endothelial-to-hematopoietic transition (EHT) [38], [39]. During EHT, a hematopoietic transcriptional program is initiated in the hemogenic endothelial cells [40]. This primes the cells for their future hematopoietic function, and in addition leads to morphological changes. After blood circulation is developed in the embryo proper, emerging hematopoietic cells migrate to the fetal liver, where EMPs proliferate and differentiate into more mature erythrocytes, megakaryocytes, and monocytes.

The third wave is known as the definitive wave of hematopoiesis and it is initiated in the aorta-gonad-mesonephros (AGM) region [33], [34]. As with the pro-definitive wave, it involves

EHT, yet at this stage the hemogenic endothelial cells give rise to hematopoietic stem cells (HSCs) with long-term ability to produce all types of blood cells. However before these emerge, hematopoietic progenitors which lack this reconstitution ability are produced. When reaching full potential as HSCs, they relocate to the fetal liver where they grow and expand, before finally homing to the bone marrow.

There are many different regulators involved in these processes and developmental stages, including essential genes *SOX17*, *GATA2*, and *RUNX1*. *SOX17* is a gene found to be imperative in promoting the endothelial cells into becoming hemogenic [34]. Research indicates that *SOX17* is not only expressed in these cells but also in emerging HSCs, and it is essential for their development [41]. Importantly, *SOX17* is also a major regulator of the endoderm germ layer during gastrulation. *RUNX1* is crucial for the differentiation of all embryonic blood cell lineages, particularly in the second and third waves of hematopoiesis, driving the formation of definitive HSPCs [35], [42]. *GATA2*, acting upstream to *RUNX1* [43], initiates the EHT in the hematopoietic program [43]. This role is crucial for specifying hemogenic endothelium and maintaining HSCs [44]. Moreover, studies point to an interplay between *SOX17* and the *RUNX1* gene where a decrease in *SOX17* expression alongside an increase of *RUNX1* marks the occurrence of the EHT [45].

As previously mentioned, studying human embryogenesis presents various challenges, including major ethical concerns (which have led to legal restrictions like the 14-day rule [11]), and difficulties related to for example the post-implantation stages, where the embryo becomes less accessible. The subject of blood development is however quite well-established in model organisms like mice, and this offers an initial, albeit limited, insight into human hematopoiesis. Advanced technologies, such as stem cell research and single-cell RNA sequencing, enable further investigation of the subject of human embryonic hematopoiesis. For instance, a study conducted in 2022 by Calvanese and colleagues [46] utilized single-cell RNA sequencing to create a transcriptome map of human hematopoietic tissues from the first trimester to birth. Their findings also highlighted HSC factors such as *RUNX1*, *HOXA9*, and *SPINK2*, where *SPINK2* is a more recently found HSC marker. Regarding studies involving stem cells, human pluripotent stem cells-derived models have proven to be of use.

### 1.1.3 Pluripotent Stem Cell-Derived Embryoid Bodies as a Model for Human Development

Via the introduction of a set of genes into adult cells, one can stimulate cells to reprogram and become pluripotent, thus generating induced pluripotent stem cells (iPSCs) [47]. Pluripotency indicates the ability to differentiate into almost any cell type, similar to the early embryonic epiblasts (see Section 1.1.2 above). iPSCs can be used to generate embryoid bodies (EBs), which are aggregates of pluripotent stem cells, that mimic some part of embryogenesis depending on the purpose of the research. This enables researchers to study various processes in embryonic development, thanks to the potential *in vitro* generated EBs have in giving rise to different cell types [48], including those involved in hematopoiesis. Studies have shown that there are different ways of enhancing hematopoiesis and generating HSCs, whereby different protocols have been established.

### Culturing Techniques and Different Protocols Utilized

Different EB culturing techniques have been developed to enhance specific lineage differentiation, such as hematopoietic differentiation. Techniques range from simple plating using ultra-low attachment plates to more complex approaches involving U-shaped or V-shaped wells that help manage cell density and aggregate size. For reproducibility and precision in cell growth and differentiation, serum-free media are often preferred so that there is better control over what proteins and supplements are added [49]. As a result of the many varied ways of culturing EBs, different protocols have been established as the research field has progressed.

In the study of hematopoiesis, particularly in relation to childhood leukemia, generating functional EBs that include cells relevant to hematopoietic development is imperative. The proper aggregation of hiPSCs is crucial for forming EBs, followed by the promotion of differentiation into specific cell types, such as mesodermal lineage cells essential for hematopoietic development. The EB generation protocol in this project adapts methodologies from two pivotal studies by Moris *et al.* (2020) [50], [51], which focus on cell aggregation and the activation of the Wnt signaling pathway to promote mesodermal lineage and embryo elongation.

The integration of these protocols aims to optimize EB formation to contribute to studies in hematopoiesis, supported by evaluating the expression of different key genes. Genes such as *RSPO3, SOX17, HOXA9, RUNX1, GATA2,* and *SPINK2*, as displayed in Figure 2. In addition to *SOX17*, the gene *CDH5* was also included in the analysis as a marker of endothelium [52], as *SOX17* is also expressed in cells of the endoderm germ layer.

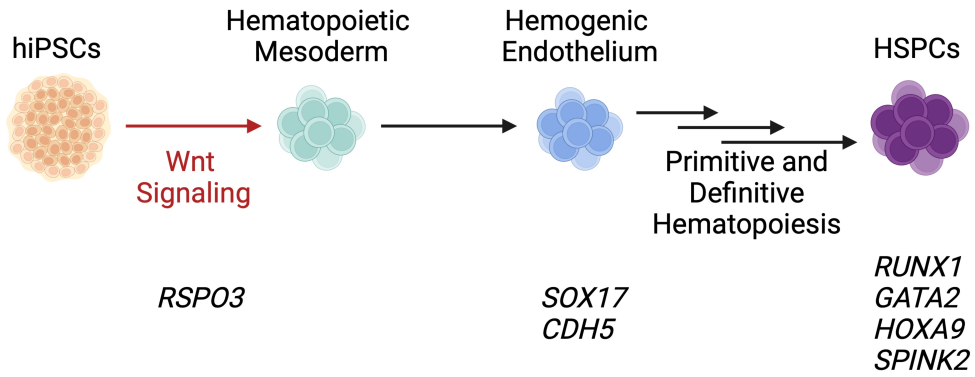


**Figure 2: Hematopoietic Development During *in vitro* Differentiation with Wnt Activation**

Overview of the development of HSPCs derived from hiPSCs enhanced by Wnt signaling. Key genes included as markers for different factors and lineages. The three parallel arrows indicate the overlapping primitive and definitive waves. Created with BioRender.com

Assessing the expression levels of these genes will enable us to track the development and to evaluate the functional integrity of EBs. The small molecule CHIR99021 (CHIR) is employed to activate Wnt signaling, marked by *RSPO3* expression. CHIR inhibits the GSK3 protein, thus hindering $\beta$-catenin degradation. Wnt activation in turn enhances HSPC development [22], [51].

## 1.2 Problem Statement and Thesis Aims

Despite the advancements in EB culturing and genetic manipulation, significant challenges remain in generating HSCs with engraftment potential [53]. The broad impact of Wnt signaling can lead to non-specific differentiation, complicating the production of targeted cell types. Moreover, the precise modulation of Wnt signaling with CHIR, crucial for optimizing differentiation, requires detailed control over timing and dosage to avoid suboptimal outcomes. Finally, achieving reproducibility in EB generation under these conditions is challenging, as slight variations in methods and conditions can lead to inconsistent results. To address these issues, this thesis will employ and evaluate the established protocol to better represent developmental processes and optimize Wnt signaling, aiming for more accurate and reproducible modeling of embryo development.

### 1.2.1 Aim and Related Questions

Building on the challenges identified in the problem statement, this section outlines the specific aims of the thesis. Each aim is designed to address a particular aspect of the challenges with generating and optimizing EBs for the study of hematopoiesis. By clearly defining these aims, the research seeks to systematically tackle the complexities of Wnt signaling modulation and its effects on cell differentiation within EBs. The following aims and related questions will guide the experimental work, aiming to enhance the reproducibility, specificity, and effectiveness of the protocols used.

**Aim 1: Confirm Wnt Pathway Activation and Optimize Modulation Conditions**

- Can Wnt pathway activation be effectively confirmed in the generated embryoid bodies, and how do different concentrations of CHIR influence this activation?

**Aim 2: Evaluate the Specificity and Efficiency of the Established Protocol**

- Does the established protocol consistently give rise to the desired cell types, specifically endothelial cells and hematopoietic stem and progenitor cells, within the generated embryoid bodies?

**Aim 3: Assess the Reproducibility of the Differentiation Protocol**

- How reproducible are the differentiation outcomes when utilizing the same established protocol with consistent CHIR concentrations across multiple experiments?

# 2

# Methods

The various methods used in this project are presented in the following sections. This includes detailing of the EB formation protocol and the subsequent analysis methods used to analyze the outcomes from the experiments where this protocol was employed. In addition, the method of single-cell RNA sequencing (RNAseq) analysis that was done on a leukemic dataset is presented. Lastly, an outline of the statistical test utilized when analyzing the significance of the real-time quantitative polymerase chain reaction (RT-qPCR) results is given.

In the initial phase of this project, samples from Experiment (Exp) IO conducted by Iman Ouadria (a previous student in the group), were analyzed. At this point in the project, RT-qPCR was employed to evaluate the gene expression profiles of EBs previously formed in Exp IO. This task served as a preliminary step, allowing me to familiarize myself with the data and the analytical techniques involved. Subsequently, I conducted my experiments involving EB formation and analysis steps, designated as Exp EHB1 and Exp EHB2. The established protocol was slightly modified after the completion of Exp IO and replicated in my subsequent experiments (See Section 2.1.3). This approach was implemented to both validate and compare the results against the results from Exp IO. The modification involves the specific method implemented when adding the CHIR treatment, which is described further in Section 2.1.3 below.

## 2.1 Embryoid Body Generation

This part involves the EB generation protocol from culturing hiPSCs, generating EBs, and harvesting samples for subsequent analysis. The protocol includes time points ranging from Day (D) -7 to D13, where D0 is the time point where the EBs are formed. The cell line used in this project is the Cellartis Human iPSC Line 22 (ChiPSC22, Takara Bio Europe AB, Gothenburg, Sweden). Unless otherwise specified, all cell culture media and other reagents mentioned in the following sections were obtained from Gibco, Thermo Fisher Scientific (Waltham, MA, USA).

### 2.1.1 Thawing and Plating of hiPSCs on Laminin Plates

Recombinant Human Laminin-521 was plated according to instructions from the manufacturer, at a coating concentration of 0.75 µg/cm$^2$. For each experiment (EHB1 and EHB2), two 6-well plates were coated with laminin, with three wells coated per plate. A total of four plates were prepared on D-7 of Exp EHB1. One plate was incubated at 37°C for two hours, and the remaining plates were stored at 4°C for up to two weeks to be used in subsequent steps in Exp EHB1 and EHB2. Two hours after plating, the laminin solution was aspirated from the wells of one 6-well plate, and 1 mL of NutriStem hPSC XF Medium (NS Medium) (Sartorius, Göttingen, Germany) was added to each of the three wells.

An aliquot of 8 mL NS Medium was prepared and heated to 37°C. A vial of hiPSCs (ChiPSC22) was removed from -140°C storage and thawed in a 37°C water bath. The thawed hiPSCs were then gently added to the NS Medium using a pipette, in a drop-wise manner to optimally separate the cells. After centrifuging at 200xg for 4 minutes, the supernatant was discarded. To achieve a concentration of 150,000 cells/mL, the cells were counted using Trypan Blue staining and a Bio-Rad automated cell counter (Hercules, CA, USA), and the necessary volume of pre-warmed NS Medium was added to the cells. The cell suspension (1 mL per well) was added to the 6-well plate, distributing the cells evenly across the surface of the dish. The cells were thereafter incubated for 48 hours at 37°C, in a 5% $CO_2$ incubator. On D-5, post-incubation, the medium was changed to fresh NS Medium after which the cells were incubated for 24 hours.

### 2.1.2 Splitting hiPSCs

On D-4, the cells were split to ensure optimal culture conditions, preventing overcrowding and maintaining the necessary environment for healthy growth and functionality. The components used in this step were pre-warmed in a 37°C water bath. These components included: Dulbecco's Phosphate Buffered Saline (DPBS) without calcium and magnesium, 2 mL per well; TrypLe Select Enzyme (TrypLe) without phenol red, 0.5 mL per well; NS Medium for neutralization and resuspension, 3 mL per well plus extra volume to achieve the desired cell concentration.

After aspirating the laminin from one of the pre-prepared 6-well plates, 1 mL of NS Medium was added to each of the three wells. The old medium in the dish with the cells was then aspirated, and DPBS was added to rinse the wells, which was subsequently aspirated. TrypLe was then added to each well, followed by incubation at 37°C for 3 minutes. The cells were pipetted up and down to dissociate them into a single-cell suspension. Then they were carefully transferred to the NS Medium for neutralization to dilute the TrypLe, ensuring that as few cells as possible were left on the old plate. After centrifuging the cells at 200xg for 4 minutes at 20°C, the supernatant was removed, and the cells were resuspended in NS Medium. Trypan Blue staining was utilized to count the cells in the Bio-Rad automated cell counter. The cell count value was then used to calculate a cell concentration of 100,000 cells/mL, and the appropriate volume of NS Medium was added to reach this concentration. Lastly, 1 mL of the cell suspension was evenly distributed to the 6-well plate already containing medium, and the plate was then incubated for 48 hours at 37°C in a 5% $CO_2$ incubator. After this incubation period, on D-2, the medium was changed to fresh NS Medium after which the cells were incubated for 24 hours under the same conditions.

### 2.1.3 Cell Aggregation and Embryoid Body Culturing

After 24 hours of incubation, on D-1, the cells were pre-treated with 3 µM CHIR by replacing the old medium with fresh NS Medium containing CHIR. The cells in one of the three wells were left untreated to serve as a control sample. This was followed by an additional 24 hours of incubation of the plate with both pre-treated and untreated cells. Importantly, CHIR was purchased as a lyophilized powder, which was reconstituted and diluted to stock concentration in dimethyl sulfoxide (DMSO) (both acquired from Sigma-Aldrich, St. Louis, MO, USA).

The hiPSC aggregation procedure to form EBs was based on the protocol published by Moris *et al.* (2020) [50]. Three aggregation media were prepared using Essential 6 (E6) Medium, differentiated by the concentration of the CHIR-treatment. This was achieved by adding a

consistent volume of CHIR (5 µL), diluted in DMSO at varied initial concentrations, to reach final concentrations of 1.8 µM, 3 µM, or 5 µM. A ROCK inhibitor (Y-27632 dihydrochloride, from Tocris Bioscience, Bio-Techne, Minneapolis, MN, USA) was also added to the media at a final concentration of 5 µM. This step involves the modification done in the protocol used in Exp EHB1 and EHB2, compared to Exp IO. The protocol implemented in Exp IO involved adding different volumes of CHIR taken from the same stock with an initial concentration of 15 mM. The protocol was then changed to use different stocks of CHIR with varied concentrations, hence adding a consistent volume of the DMSO carrier to the media instead.

On D0, the pre-treated cells were aggregated to initiate EB formation using the following procedure. From the 6-well plate, the old medium was aspirated, and the wells were rinsed with 2 mL of DPBS per well, which was then aspirated. TrypLe (0.5 mL per well) was added, followed by incubation for 3 minutes at 37°C. The cells were pipetted up and down to dissociate them into a single-cell suspension, before being collected into 4 mL of DPBS. This was followed by centrifugation at 200xg for 4 minutes. The supernatant was removed, and the cells were resuspended in 1 mL E6 Medium (without CHIR) for counting. Both the pre-treated cells and the untreated cells were counted, as previously described. The volume of cell suspension added to each of the aggregation media (with CHIR) was calculated to achieve a concentration of 15 cells/µL. Thereafter, 40 µL of this media was pipetted into each well of a U-bottom 96-well plate, positioning the droplets at the bottom of the well. This resulted in around 600 cells per well. Five and a half 96-well plates were prepared per CHIR condition for RT-qPCR analysis. Finally, the plates were centrifuged at 700 rpm for 2 minutes to aggregate the cells at the bottom of the wells, and then they were incubated for 24 hours at 37°C in a 5% $CO_2$ incubator.

The days following D0 involved specific media changes in all the wells of the 96-well plates, with each volume of media addition or removal being consistently 150 µL. These operations were performed from the side of the wells to avoid disturbing the EBs. Between these steps, the plates were incubated at 37°C in a 5% $CO_2$ incubator. Fresh Essential 6 (E6) Medium was added on D1, and on D2, the old medium was replaced with an equal volume of fresh E6 Medium. At the time points that followed, up until D13, the protocol for EB culturing was exchanged for the protocol published by Ditadi and Sturgeon (2016) [54] with some alterations. The alterations were made to the D3 medium according to Oburoglu *et al.* (2022) [55], and to the D8 medium where vascular endothelial growth factor (VEGF) and basic fibroblast growth factor (bFGF) were supplemented. For details about the media with supplements utilized in the following steps, see Table A.1. On D3, the media was changed, this time replacing the old medium with fresh StemPro-34 SFX (SP34)-d3 Medium, which contained 3 µM CHIR and 1 ng/mL Activin A. On D4, a change to fresh SP34-d4 Medium was conducted. After 48 hours, on D6, the media was again changed, this time to SP34-d6 Medium. On D8 and D10, the medium was replaced with SP34-d8 Medium containing VEGF.

### 2.1.4 Harvesting Cells for RT-qPCR

Cells and EBs were harvested during the embryoid body generation Exp EHB1 and EHB2, at the following time points: D0, D3, D8, and D13. On D0, two samples were collected per experiment. One sample consisted of untreated cells and the other sample consisted of pre-treated cells, referred to in Section 3 as before pre-treatment (BPT) and D0, respectively. On Days 3, 8, and 13, two replicate samples of EBs per CHIR condition were collected, consisting of independent pools of EBs. This equaled a total of six samples, including duplicates for each of the time points D3, D8 and, D13 (see Figure 3d).

BPT and D0 cells were harvested by collecting the medium containing the cells dissociated with TrypLe (dissociation step as described in Section 2.1.3 above) into separate tubes. EBs were harvested by the collection of EBs from the 96-well plates for each condition into separate tubes. Then centrifugation at 200xg for 4 minutes was done. The supernatant was aspirated, and the pellet was rinsed with DPBS, for each sample. Following another round of centrifugation and aspiration of the supernatant, the pellet was resuspended in RLT Buffer. The samples were then vortexed and stored at -80°C.

## 2.2 RNA Extraction, cDNA Synthesis, and RT-qPCR

Total RNA was extracted using either the RNeasy Micro Kit or the RNeasy Mini Kit (both acquired from Qiagen, Valencia, CA, USA), depending on the amount of cell material in the samples. The micro kit was used for D3 and D8 samples, while the mini kit was used for BPT, D0, and D13 samples. Complementary DNA (cDNA) synthesis was conducted by reverse transcribing the RNA with either QuantiTect Reverse Transcription Kit (Qiagen, Valencia, CA, USA) or the SuperScript III First-Strand Synthesis System with random hexamer primers (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), according to the manufacturers' instructions.

RT-qPCR was performed with the SsoFast EvaGreen Supermix (Bio-Rad, Hercules, CA, USA) using a 7900 HT Fast Real-Time PCR System in a 96-well plate format (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA). Primers for the RT-qPCR were designed using the NCBI Primer-BLAST tool to ensure specificity to the target sequences, with primer efficiency validated prior to experiments. All primers except for the ones used to evaluate *CDH5* expression were designed by Iman Ouadria. The ones for *CDH5* were designed by me. Primer sequences and details are summarized in Table A.2.

Each RT-qPCR reaction mixture consisted of 5 µL of SsoFast EvaGreen Supermix, 0.5 µL of each primer (10 µM), 2.5 µL of cDNA template, and 1.5 µL of nuclease-free water to a final volume of 10 µL. The *SOX17* primers were added in half volume, with the other half as nuclease-free water for optimal amplification of this gene. Each sample was analyzed in triplicate in the RT-qPCR. The cycling conditions were as follows: 50°C for 2 minutes, 95°C for 10 minutes, 40 cycles of denaturation at 95°C for 15 seconds, and annealing/extension at 60°C for 1 minute. A melting curve analysis was conducted post-amplification to verify the specificity of the PCR products. The protocol included: 95°C for 15 seconds, 60°C for 15 seconds, and lastly 95°C for 15 seconds.

### Optimization of the Primer Pairs Utilized in the RT-qPCR

Primer pairs designed prior to the start of this project were optimized by Iman Ouadria. Optimization was done via RT-qPCR with hiPSC cDNA dilution series, and primer efficiencies were calculated by the standard curve method (see Table A.2 for efficiencies). The optimization also included dissociation curve evaluation for each primer, to ensure specific amplification. *HOXA9* primer pair efficiency could not be determined. However, by evaluating the dissociation curve during optimization and also during the RT-qPCR analysis of the samples, specific amplification was ensured.

For *RSPO3* and *CDH5*, some unspecific amplification was observed. Due to this, the primers for these genes were further optimized by doing RT-qPCR on a cDNA dilution series using cDNA

from hiPSCs, followed by a calculation of primer efficiency by the standard curve method (see Figure S1). The results from this displayed that unspecific amplification occurred when the cDNA content was too low. This was taken into account in the continued RT-qPCR analysis, by re-running reactions that displayed unspecific amplification using a reduced cDNA dilution factor.

### 2.2.1 RT-qPCR Data Handling and Gene Expression Quantification

RStudio (v.4.3.1, R Foundation for Statistical Computing) was used to sort the data produced from the RT-qPCR analysis. This was done by removing data unnecessary to the analysis, ordering the data in a uniform way to make the analysis more efficient, and removing any outliers from the amplification. Outliers were identified by evaluating the amplification and dissociation plots from the RT-qPCR. When unspecific amplification or otherwise diverging data was observed, these were excluded from the analysis (see Figure S2 for representative dissociation curves for both specific and unspecific amplification). In addition to the data sorting, RStudio was also used to calculate the average value of the RT-qPCR triplicate data for subsequent analysis steps.

The code used for both sorting and analyzing data can be viewed via the following link: https://github.com/ElinHilpoldBerntsson/MasterThesis_ReferenceCode.git. This code was specifically used to sort and analyze Exp IO data. Here it is used as a reference code for all experiments, as the code used in Exp EHB1 and EHB2 analysis was of similar structure, only with some changes to fit the data optimally.

Relative gene expression levels were quantified using the $2^{-\Delta Ct}$ method, normalized to the expression of housekeeping genes. Housekeeping genes used were glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and $\beta$-actin (ACTB). An average of the Ct values of the housekeeping genes was used for a more reliable and robust analysis. Fold changes in gene expression were calculated using the $2^{-\Delta\Delta Ct}$ (Livak) method, where values between experimental and control groups were compared. The formulas utilized to calculate both relative expression and fold change are presented below.

#### The $2^{-\Delta Ct}$ Method

In order to calculate the relative expression of each gene included in the RT-qPCR, the following formulas were used. First the $\Delta$Ct is calculated, followed by an exponentiation by 2 to get the relative expression.

$$\Delta Ct = Ct\ (target) - Ct\ (reference) \tag{2.1}$$

$$\text{Relative expression} = 2^{-\Delta Ct} \tag{2.2}$$

Here, target stands for target gene and reference represents housekeeping gene.

#### The Livak Method: A Simplified Version of the Pfaffl Method

The Pfaffl method offers a precise approach to RT-qPCR fold change quantification by accounting for variations in the efficiency across different primers [56]. It employs the actual efficiencies of primers for both target and reference genes, utilizing the following formula:

$$\text{Fold change} = \frac{(E_{\text{target}})^{\Delta Ct(\text{untreated}-\text{treated})}}{(E_{\text{reference}})^{\Delta Ct(\text{untreated}-\text{treated})}} \tag{2.3}$$

Where untreated and treated signifies untreated sample and treated sample, respectively. This calculation provides an accurate measurement of gene expression levels by adjusting for the specific amplification efficiencies, which is essential for reliable quantification.

In the present report the Livak method was employed, which simplifies the analysis by assuming optimal primer efficiencies, standardized at an efficiency of 2. This results in the formula $2^{-\Delta\Delta Ct}$, simplified as follows:

$$\text{Fold change} = \frac{2^{\Delta Ct,\text{target(untreated - treated)}}}{2^{\Delta Ct,\text{reference(untreated - treated)}}} \tag{2.4}$$

$$= 2^{-((Ct,\text{target(treated)}-Ct,\text{target(untreated)})-(Ct,\text{reference(treated)}-Ct,\text{reference(untreated)}))} \tag{2.5}$$

$$= 2^{-\Delta\Delta Ct} \tag{2.6}$$

where $\Delta\Delta$Ct is the difference in $\Delta$Ct values between treated and untreated samples, noted as the fold change. $\Delta$Ct is the threshold cycle difference between the target and reference genes, as a measurement of relative expression as detailed above.

## 2.3 Flow Cytometry

The EBs were harvested on D8 and D13 for flow cytometry analysis, by transferring them from the 96-well plates into 15 mL conical tubes. Each well was washed with DPBS, and the wash was added to the corresponding tubes. After settling, the supernatant was transferred to new 15 mL tubes equipped with mesh (Filcon strainers). The original tubes were washed again with DPBS, which was then also transferred to the mesh tubes.

Dissociation protocols differed between the two harvest times. At D8, 0.5 mL of TrypLe was added to each tube. The tubes were gently tapped to mix and then incubated at 37°C for 3 minutes. This mixing and incubation cycle was repeated three times until complete dissociation of the EBs. Following dissociation, the contents were washed with DPBS and transferred to the mesh tubes. At D13, 1 mL of Liberase TM solution was added directly to the EB tubes, followed by incubation at 37°C for 20 minutes, with intermittent tapping to mix. Mechanical fragmentation was facilitated by gentle pipetting. After this, the supernatant was transferred to the mesh tubes. The remaining EBs were suspended in TrypLe, incubated, and tapped as described for D8. After three cycles, the supernatant was again transferred to the mesh tubes, leaving a residual volume for mechanical dissociation via gentle pipetting. Tubes were then washed with DPBS, and all contents were transferred to the mesh tubes.

After the dissociation, the tubes were centrifuged at 200xg for 7 minutes, and the supernatant was aspirated. Cells were resuspended in 100 µL of FACS buffer (DPBS containing 2% fetal bovine serum). Cell counts were performed using Trypan Blue with the Bio-Rad counter, noting the real volume of each sample. Cells were then resuspended in 1-3 mL of FACS buffer and

centrifuged at 300xg for 5 minutes. The supernatant was discarded, and cells were treated with the block master mix for subsequent analysis.

Before sample staining and flow cytometry were conducted, different components were prepared. The single stains were prepared by adding a drop of pre-mixed beads to 100 µl of FACS buffer in each bead tube, excluding the Fc block. Then the antibodies were added, see Table A.3 for specifics regarding these. For the Fluorescence Minus One (FMO) controls, 20-30% of the remaining cells were pooled, estimating about 10,000 cells per antibody. A master mix for the antibodies and Fc block was also prepared. To each sample and FMO, 30 µl of Fc block master mix was added. For the unstained controls, 300 µl of FACS buffer, containing double the cells relative to the FMOs, was added. 50 µl of FACS buffer was added to the FMOs along with any prepared FMO mixes. Antibodies were introduced into the antibody master mix, and FMOs as needed. Finally, 50 µl of the antibody master mix was distributed into each sample.

The samples were stained for 30 minutes or more. 3 mL of FACS buffer was added to each stained tube and then the tubes were centrifuged at 200xg for 4 minutes. The supernatant was aspirated carefully and then 200 µL FACS buffer was added to the tubes. Thereafter the flow cytometry was conducted utilizing the BD LSRFortessa with the BD FACSDiva Software (BD Biosciences, Franklin Lakes, NJ, USA). The flow cytometry data was analyzed, and the gating strategy was done in FlowJo (v.10.10, BD Biosciences).

## 2.4 Single-cell RNAseq Analysis of a Leukemic Dataset

To analyze gene expression within leukemic cells, a Uniform Manifold Approximation and Projection (UMAP) plot was generated using a comprehensive single-cell RNA-seq dataset from a transgenic mouse model of AML with an MLL translocation [57]. This mouse model expressed a human *EIF6* transgene by a doxycycline (DOX)-responsive promoter, and the dataset includes 8 samples where four are under DOX condition and four are of NO-DOX condition.

The UMAP was plotted using the ScarfWeb platform (Nygen Analytics), an online tool tailored for analyzing single-cell RNAseq data. Initially, cells of low quality were excluded based on different parameters, such as transcripts per cell (Figure S3). This was followed by the identification of highly variable genes, which contribute most to the variability between cells (Figure S4). Although different batch corrections were considered to address technical variability, it was ultimately decided to proceed without applying any batch correction.

The UMAP method visualizes single-cell data by reducing its dimensionality, where each point represents a cell. Cells with similar gene expression profiles tend to cluster together, indicating similarity. UMAPs effectively highlight local affinities, especially where clusters are directly adjacent to each other. However, broader interpretations should be approached with caution, as distance in a UMAP space does not necessarily correlate with similarity.

## 2.5 Data Plotting and Statistical Analysis

The plots displaying RT-qPCR and flow cytometry data in this thesis were created with the GraphPad Prism 10 Software (v.10.2.2, Dotmatics). The results shown represent mean ± standard deviation (SD). Two-way Analysis of Variance (ANOVA) was employed as a statistical test of the RT-qPCR and flow cytometry data. Here, the Tukey multiple comparisons test was

utilized. The results from the statistical test were included on the bar plots. Asterisks indicate levels of statistical significance (*p≤0.05, **p≤0.01, ***p≤0.001, ****p≤0.0001).

# 3

# Results

This section presents the results obtained from the various analytical methods utilized in this project. Initially, an overview of the experimental protocol is provided, detailing the sample distribution, relevant gene markers, and their hypothesized expression patterns within the generated EBs. Subsequent sections go further into the gene expression results, discussing each gene of interest in the context of its role in the specific signaling pathway or lineage determination.

To put the gene expression analyzed in this project into a broader context regarding embryonic developmental processes, a UMAP from Pijuan-Sala et al. [58] is utilized. This UMAP visualizes cellular differentiation from pluripotent state in mouse embryos from E6.5 to E8.5. It captures the dynamic changes and lineage specification of pluripotent epiblast cells as they diverge into ectodermal, mesodermal, and endodermal lineages, encompassing the key phases of gastrulation and early organogenesis.

Additionally, the UMAP generated from a comprehensive single-cell RNA-seq dataset derived from AML cells (Section 2.4) is included to illustrate the expression patterns of the genes of interest within the leukemic cellular landscape.

The UMAP visualizations are repeatedly presented in the following sections, with specific genes plotted to highlight their potential roles and interactions in the broader contexts mentioned. This serves to bridge the gene expression analyses with detailed datasets, enhancing the understanding of the genes' functions in varied biological settings, as described more in depth in Section 1.1.2.

## 3.1 Wnt Signaling: Impact on Cell Morphology and Lineage Emergence in Embryos

To investigate the effect of Wnt activation on the emergence of desired cell lineages in hiPSC-derived EBs, an established experimental protocol was implemented as described in Section 2.1. Figure 3a outlines the timeline and major steps of this protocol, from thawing and plating hiPSCs to the downstream steps such as treatment with CHIR, EB formation, cell harvesting, and subsequent RT-qPCR analysis. This full workflow was replicated in Exp EHB1 and EHB2 to assess reproducibility.

The morphology and the progress of culturing of the hiPSCs were evaluated during the period leading up to EB formation (Figure 3b). The hiPSCs prior to the pre-treatment exhibited elongated and pointy morphologies, characterized by sharp, angular cell borders and extended projections. This phenotype is typical for undifferentiated stem cells before the influence of lineage-specifying signals takes place [59]. Following the application of CHIR, a notable tran-

sition in cell morphology was observed (D0 PT). The cells appeared less elongated and the previously sharp projections became somewhat subdued, as the cellular outlines became more rounded. This morphological change is consistent with the onset of differentiation processes, likely driven by the activation of the Wnt signaling pathway via CHIR treatment.

On D0 in the EB formation phase, cells were treated with CHIR at concentrations of 1.8 µM, 3 µM, and 5 µM, to evaluate the effect of Wnt signaling. The growth patterns following EB formation under these conditions are documented in Figures 3c and S5. Microscopic analysis at various time points showed that the EBs in all three conditions displayed significant growth and compactness. However, no pronounced differences in elongation or overall size were observed between the EBs treated with 1.8 µM and 3 µM CHIR. Notably, the EBs exposed to 5 µM CHIR initially followed a similar morphological pattern but diverged by D13, presenting as somewhat smaller and with signs of cells dispersing from the main body of the EB. These dispersed cells potentially represent differentiated cell types, possibly blood cells, suggesting a differential response to the higher concentration of CHIR.

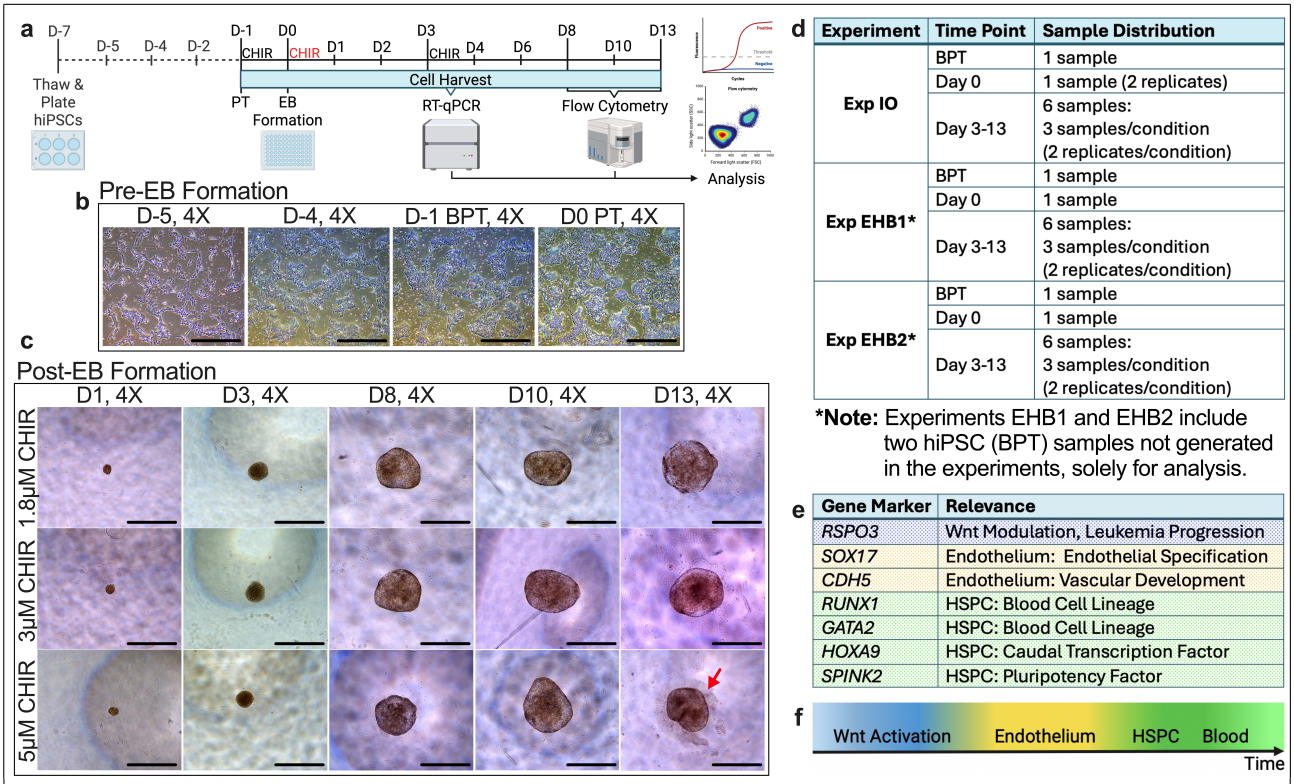


**Figure 3: Experimental Workflow, Key Analysis Steps, and Sample Outline**

**a,** Schematic illustration of the workflow and timeline for EB generation, cell harvesting and subsequent analysis steps. CHIR was added on three separate time points as indicated in the Figure. The text in black represents an addition of 3 µM of CHIR, while the text in red represents and addition of 1.8, 3, or 5 µM of CHIR as experimental conditions. Created with BioRender.com **b,** Representative microscopy images at 4X magnification of hiPSCs shown at three days before pre-treatment (D-5, D-4, D-1) and after pre-treatment (D0), all preceding EB formation. **c,** Representative microscopy images at 4X magnification of developing hiPSC-derived EBs at Days 1, 3, 8, 10, and 13 post-EB formation (notated as D1, D3, D8, D10, and D13).

Dispersed cells indicated by the red arrow. Note: Scale bars represent 1 mm. **d,** Table of the sampling framework for Experiments IO, EHB1 and EHB2, detailing the number of samples, including replicates, collected. See also Table A.4 **e,** Table of RT-qPCR gene markers and their relevance to the study. **f,** Timeline of the expected emergence of Wnt activation and cell lineages. Created with BioRender.com

Throughout the experimental protocol, the produced samples captured the state of the cells and/or EBs at critical time points, specifically before BPT, and on D: 0, 3, 8, and 13. The distribution of samples and replicates collected at these time points across all three experiments (Exp IO, EHB1, and EHB2) are detailed in Figure 3d. Samples from Exp IO, along with those produced in Exp EHB1 and EHB2, were all used in RT-qPCR analysis to assess gene expression profiles. To enhance the robustness of the dataset, particularly for untreated hiPSCs, two additional BPT samples generated in other experiments were also analyzed by RT-qPCR and included in the data analysis of Exp EHB1 and EHB2 datasets. For more details about the samples, see Table A.4.

The emergence of specific cell lineages in the EBs was evaluated by amplifying targeted gene markers via the RT-qPCR analysis. The primary focus was on genes that are central in the activation of the Wnt signaling pathway and the differentiation into endothelial and HSPCs. Specifically, *RSPO3* was analyzed for its role in Wnt signaling; *SOX17* and *CDH5* were assessed for the emergence of endothelial lineage; and *RUNX1, GATA2, HOXA9,* and *SPINK2* were examined for their involvement in HSPC development (Figure 3e). These genes not only serve as markers for these overarching factors but also play significant roles in more specialized cellular functions, which are described in more depth in Section 1.1.2.

The initial hypothesis regarding pathway activation and specific lineage emergence was formulated during the analysis of RT-qPCR results using samples from Exp IO. This analysis provided foundational insights into gene expression patterns that informed subsequent adjustments to the analytical methods, such as the introduction of an additional specific marker (*CDH5*) in Exp EHB1 and EHB2. This aimed to refine the understanding of endothelium lineage emergence, ensuring more accurate conclusions concerning this aspect.

The RT-qPCR analysis was expected to show distinct patterns of gene expression over time, indicating specific lineage emergence as demonstrated in the timeline depicted in Figure 3f. For instance, an early and notable increase in *RSPO3* expression was anticipated, indicative of Wnt pathway activation by the addition of CHIR. *SOX17* and *CDH5* were thought to have an evident increase in expression following this, suggesting the presence of endothelium lineage, thereafter potentially leading to HSPC emergence via the endothelial-to-hematopoietic transition (EHT).

The gene expression results from RT-qPCR analysis for all markers are presented in the subsequent sections, primarily displayed as fold change relative to the hiPSC control group (BPT samples). Specifically, results from Exp IO are presented in parallel to the combined results from Exp EHB1 and EHB2, enabling a comparative analysis. Additionally, pooled data from all three experiments is presented with statistical testing to provide a comprehensive overview of the observed patterns.

## 3.2 *RSPO3* Expression Indicates Wnt Signaling Activation and Lineage Specification in Embryoid Bodies

The expression of *RSPO3*, as anticipated, showed an upregulation following CHIR addition at D0, which activates the Wnt signaling pathway closely associated with this gene (Figures 4a, b). This increase was more evident in Exp EHB1 and EHB2. Notably, the response to CHIR at D3 in Exp IO was distinct, showing a more pronounced increase in expression at higher CHIR concentrations compared to the relatively uniform response across conditions in Exp EHB1 and EHB2. In Exp IO, *RSPO3* expression for the two lower CHIR concentrations continued to increase until D8, peaking at this time point. In contrast, in Exp EHB1 and EHB2, *RSPO3* expression peaked earlier at D3. After reaching their respective peaks, the expression levels either stabilized or slightly declined in all experiments. However for the 5 µM condition, a diverging pattern was displayed in all three experiments, with a significant decline at D13 (Figure 4c).

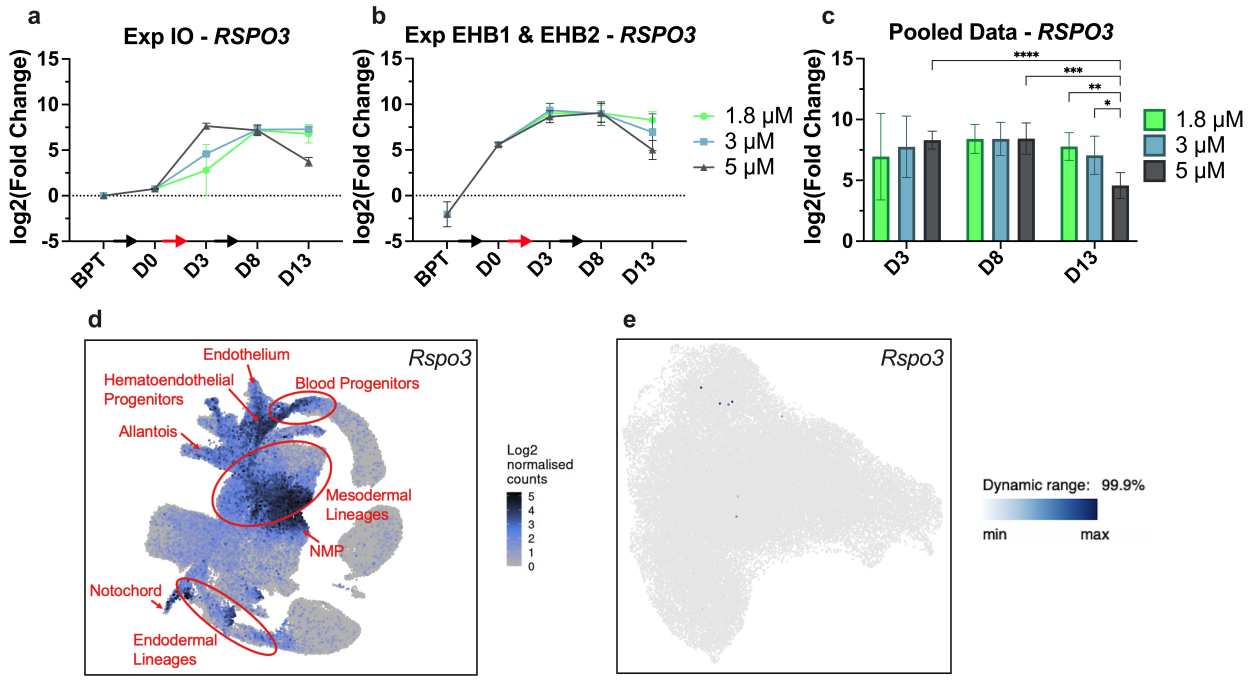


**Figure 4: *RSPO3* Expression in Experimental Data and Lineage Associations**

**a,** *RSPO3* expression measured by RT-qPCR in Exp IO over time. BPT with n=1, and remaining time points with n=2 for each. **b,** *RSPO3* expression in Exp EHB1 and EHB2 over time. D0 with n=2, and remaining time points with n=4 (2 independent experiment) for each. Black arrow: 3 µM CHIR. Red arrow: 1.8, 3 or 5 µM CHIR. **c,** Combined *RSPO3* expression from Exp IO, EHB1, and EHB2, with n=6 per time point. Plots display $\log2(2^{-\Delta\Delta Ct})$. Values represent mean ± SD. Asterisks indicate levels of statistical significance as determined by two-way ANOVA and Tukey multiple comparisons test (*p≤0.05, **p≤0.01, ***p≤0.001, ****p≤0.0001). **d,** UMAP of single-cell RNAseq data of mouse gastrulation and early organogenesis from [58], colored by *Rspo3* expression level. Embryonic cell populations with notable *Rspo3* expression are outlined in red. NMP: Neuromesodermal progenitors. See also Fig. S6. **e,** UMAP of single-cell RNAseq data of a mouse model of MLL-fusion leukemia from [57], colored by *Rspo3* expression level. See also Fig. S7a.

The UMAP shown in Figure 4d illustrates significant *Rspo3* expression in several key cell lineages, particularly those associated with the mesoderm, such as blood progenitors, hematoendothelial progenitors, and endothelium. It also indicates *Rspo3*'s involvement in the endodermal lineage. The observed increase in *RSPO3* expression (Figures 4a-c), where CHIR was used to activate Wnt signaling, suggests a potential enhancement or initiation of differentiation processes within these highlighted lineages. Given the significant role of Wnt signaling in the development and differentiation of endothelial and HSPC lineages, an elevation in *RSPO3* expression could indicate that these processes are being actively promoted within the generated EBs.

Interestingly, in contrast to the *Rspo3* expression in the UMAP depicting gastrulation and early organogenesis, the UMAP based on the leukemic dataset shows an almost nonexistent *Rspo3* expression (Figure 4e). Wnt signaling pathways are crucial in various cancers, including AML, for maintaining the stemness and survival of cancer cells, especially leukemia stem cells (LSCs) [60], [61].

In spite of the near absence of *Rspo3*, prominent expression levels of other Wnt pathway markers such as *Ctnnb1* ($\beta$-catenin) and *c-Myc* (Figures S7d, e), suggest that Wnt pathway is activated. However, *Rspo3* might be expressed in niche cells not included in this dataset.

These results indicate an activation of the Wnt signaling pathway in the generated EBs, although it may not necessarily be sensitive to a higher CHIR concentration (Figures 4a-c). As demonstrated by Figure 4d, the *Rspo3* expression is prominent in crucial cell lineages related to hematopoiesis. This points to the role Wnt signaling plays in initiating hematopoietic lineages.

## 3.3 Endothelial Lineage Emergence in Embryoid Bodies

The expression of *SOX17* increased after pre-treatment with CHIR, with a more notable increase observed in Exp EHB1 and EHB2 at D0, compared to Exp IO (Figures 5a, b). In Exp IO, *SOX17* expression varied more across CHIR conditions, whereas in Exp EHB1 and EHB2, the expression followed a more uniform pattern. This suggests different effects of CHIR between these experiments. Across all experiments, *SOX17* expression significantly peaked at D8, followed by a decrease, particularly in the 5 µM CHIR condition (Figure 5d).

*CDH5* expression in Exp EHB1 and EHB2 followed a similar pattern to *SOX17*, with an initial increase post D0 and a significant peak at D8. The highest expression decline was observed in the 5 µM condition by D13, as shown in Figure 5e.

The UMAPs (Figures 5g, h) highlight the cell lineages associated with *Sox17* and *Cdh5* expression in embryonic development. While *Sox17* serves as an endothelium marker, the UMAP confirms its additional involvement in various endodermal lineages, as detailed in Section 1.1.2. The results of *CDH5*, analyzed together with *SOX17* in Exp EHB1 and EHB2, verify the emergence of endothelium lineage in the EBs. Although similar assumptions might be made for Exp IO, the absence of *CDH5* in the analysis of these samples prevents a definitive confirmation from being made.

When analyzing the expression of *Sox17* in the leukemia dataset, no significant data was available to map. The UMAP in Figure 5f shows an almost absent expression of *Cdh5*. These results are in line with the dataset only comprising blood cells, hence the lack of evident endothelium expression. Nevertheless, given that *Cdh5* also marks a subset of HSCs [62], the expression

observed in this map may reflect the presence of these cells.

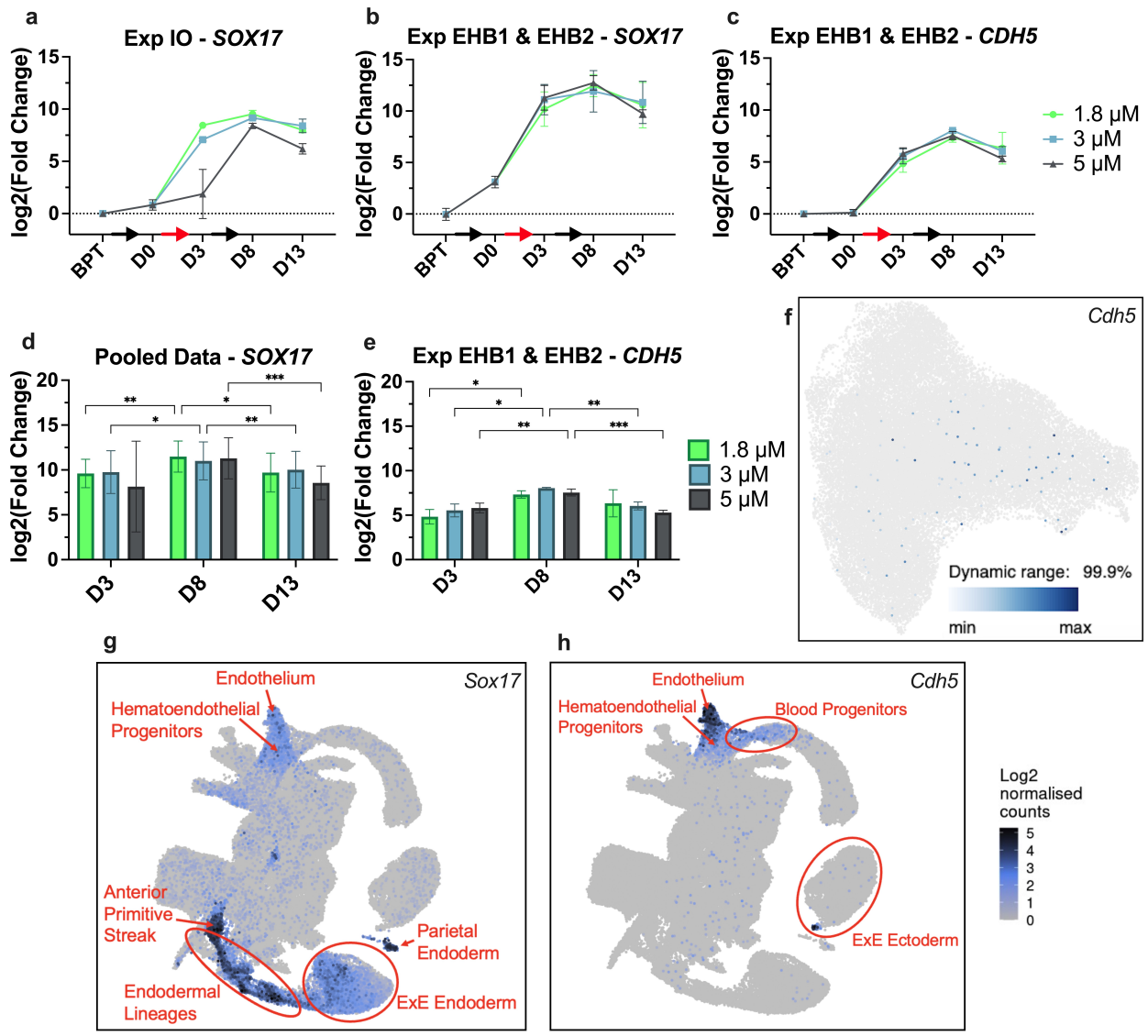


**Figure 5: Expression Profiles of Endothelial Markers**

**a-e**, RT-qPCR analysis of *SOX17* (**a, b, d**) and *CDH5* (**c, e**) over time in Exp IO, EHB1, and EHB2. **a,** *SOX17* in Exp IO, with n=2 for each time point except for BPT (n=1). **b, c,** Expression of *SOX17* (**b**) and *CDH5* (**c**) in Exp EHB1 and EHB2, with n=4 (2 independent experiments) for each time point, except for D0 (n=2). Black arrow: 3 µM CHIR. Red arrow: 1.8, 3 or 5 µM CHIR. **d**, Pooled data for *SOX17* expression from all three experiments, with n=6 per time point. **e,** *CDH5* expression in Exp EHB1 and EHB2 presented as a bar plot, n=4 (2 independent experiments) per time point. Plots display $\log2(2^{-\Delta\Delta Ct})$. Values represent mean ± SD. Asterisks indicate levels of statistical significance as determined by two-way ANOVA and Tukey multiple comparisons test (*p≤0.05, **p≤0.01, ***p≤0.001). **f,** UMAP of single-cell RNAseq data of a mouse model of MLL-fusion leukemia from [57], colored by *Cdh5* expression level. See also Fig. S7a. **g, h,** UMAPs of single-cell RNAseq data of mouse gastrulation and early organogenesis from [58], colored by *Sox17* and *Cdh5* expression level, respectively. Embryonic cell populations with notable gene expression are outlined in red. ExE: Extra-embryonic. See also Fig. S6.

The results presented highlight the significant associations of *SOX17* and *CDH5* in endothelial cell differentiation within EBs. As described in Section 1.1.2, endothelial cells play a crucial role in the EHT. The emergence of the endothelium lineage, as indicated by these results, suggests that the development of HSPCs might have been promoted, potentially giving rise to the blood cell lineage.

## 3.4 Complementary Dynamic of Markers for Blood Lineage Specification

The expression of *RUNX1* in Exp EHB1 and EHB2 followed a similar trend to that observed in Exp IO, with an initial increase following time point BPT, a decrease at D3, and a subsequent rise at D8 (see Figures 6a, b). Notably, the increase from D3 to D8 was more pronounced in Exp EHB1 and EHB2 compared to Exp IO. Moreover, all conditions demonstrated an upregulation at D13, unlike in Exp IO. This variation contributed to significant variability in the pooled data at D13, as shown in Figure 6c. Significant changes in *RUNX1* expression were observed at D8 and D13 relative to D3 in the combined data, indicating enhanced activity of this gene as the generation of the EBs progressed.

Concerning *GATA2*, divergent expression patterns were observed in Exp IO across conditions, compared to the more consistent trends in Exp EHB1 and EHB2 (Figures 6d, e). In Exp IO, lower CHIR concentrations led to an initial increase in expression after D3, whereas in the 5 µM condition, the increase occurred earlier. Post-D8, expression stabilized in the 3 µM condition and declined in the higher concentration settings. In contrast, Exp EHB1 and EHB2 showed a continuous increase in *GATA2* expression, peaking at D13 (Figure 6e). The combined data shown in Figure 6f reveals substantial variability in expression levels at D3; however, a significant rise was observed between D3 and D13 for the lower CHIR concentrations.

As outlined in Section 1.1.2, there is an interlinked relationship between *RUNX1* and *GATA2* in the hematopoietic development. This connection can be seen in Exp IO, where the expression patterns of both genes exhibit similar trends at D0 and forward (Figures 6a, d). However, this pattern is less evident in Exp EHB1 and EHB2. In these experiments, the expression of *RUNX1* and *GATA2* displays a potentially complementary rather than parallel relationship. When one gene increases, the other either stabilizes or is downregulated. This interaction confirms previous findings regarding these genes, where downregulation of *RUNX1* leads to upregulation of *GATA2* [43]. Nevertheless, both genes show statistically significant upregulation over time in the lower CHIR conditions (Figures 6c, f).

The similarities between *Runx1* and *Gata2*, regarding their expression in different cell lineages, are displayed in Figures 6g, h. When mapping the gene expressions in the UMAP representing gastrulation and early embryogenesis, both showed expression in blood progenitors and erythroid lineages. This underlines their relevance as markers for HSPCs, specifically blood cell lineage, as presented in Section 1.1.2. Both genes also displayed expression in for instance endoderm lineages, indicating a broader involvement in the embryogenesis. Notably, *Gata2* expression was also indicated in the endothelium and hematoendothelial progenitors. This aligns with its role in the EHT, where it acts as an early marker in the emergence of HSCs.

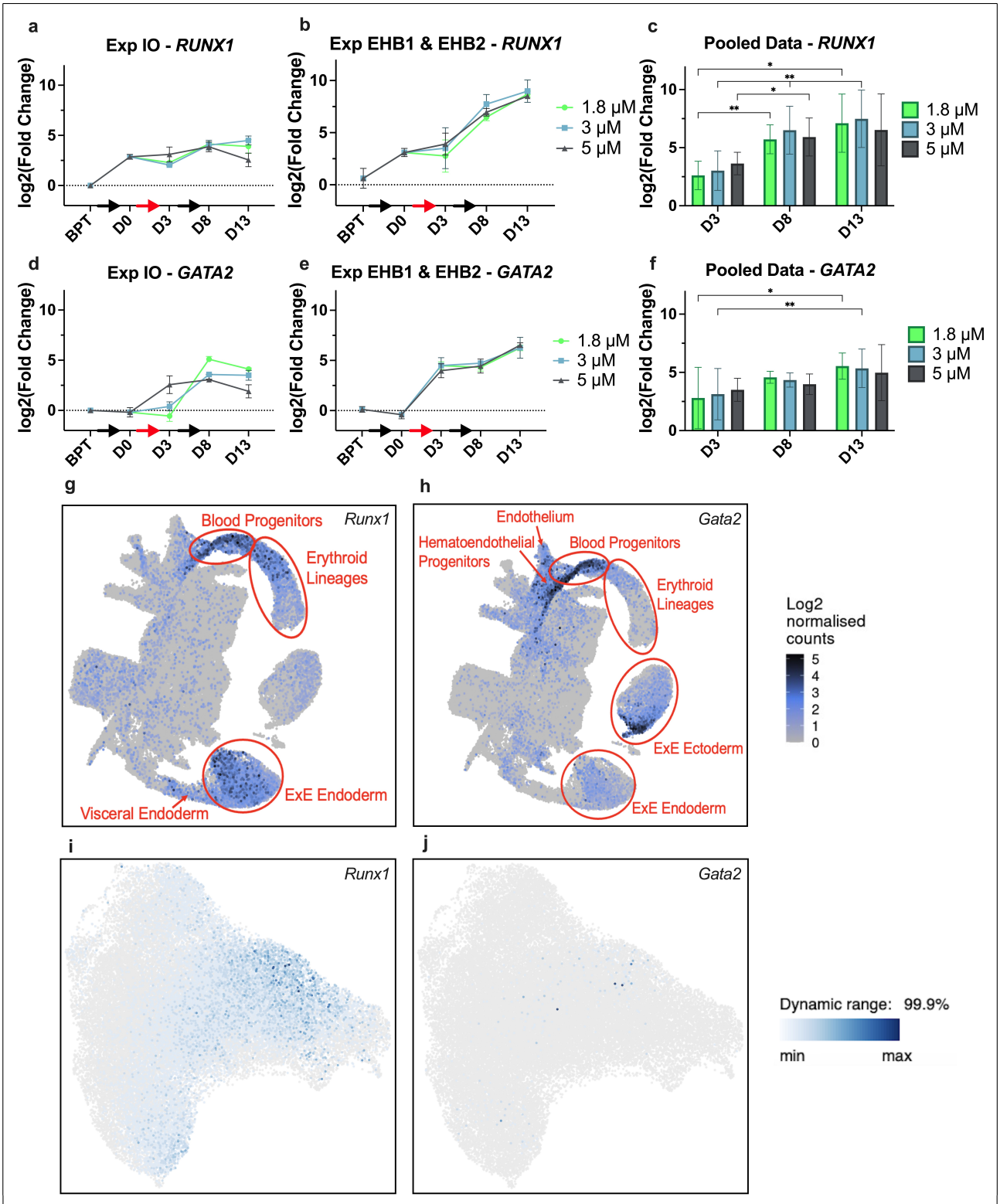


**Figure 6: Expression Profiles of Blood Lineage Markers**

**a-f,** RT-qPCR analysis of *RUNX1* (**a, b, c**) and *GATA2* (**d, e, f**) over time in Exp IO, EHB1, and EHB2. **a, d,** Expression of *RUNX1* (**a**) and *GATA2* (**d**) in Exp IO, with n=2 for each time point except for BPT (n=1). **b, e,** Expression of *RUNX1* (**b**) and *GATA2* (**e**) in Exp EHB1 and EHB2, with n=4 (2 independent experiments) for each time point, except for D0 (n=2). Black arrow: 3 µM CHIR. Red arrow: 1.8, 3 or 5 µM CHIR. **c, f,**

Pooled data for *RUNX1* expression (**c**) and *GATA2* expression (**f**) from all three experiments, with n=6 per time point. Plots display $\log2(2^{-\Delta\Delta Ct})$. Values represent mean ± SD. Asterisks indicate levels of statistical significance as determined by two-way ANOVA and Tukey multiple comparisons test (*p≤0.05, **p≤0.01). **g, h,** UMAPs of single-cell RNAseq data of mouse gastrulation and early organogenesis from [58], colored by *Runx1* and *Gata2* expression levels, respectively. Embryonic cell populations with notable gene expression are outlined in red. ExE: Extra-embryonic. See also Fig. S6. **i, j,** UMAPs of single-cell RNAseq data of a mouse model of MLL-fusion leukemia from [57], colored by *Runx1* and *Gata2* expression levels, respectively. See also Fig. S7a.

*Runx1* expression was observed in the leukemia dataset, most evident in clusters representing more mature hematopoietic cells, as shown in Figures 6i and S7b. The expectation of *Gata2* expression in the less mature cells, as it is an early HSC marker, did not prove right. As displayed by Figure 6j, *Gata2* did not exhibit notable expression in these leukemia samples. Keeping in mind that the analysis is done on leukemia cells, this might have had an impact in the *Gata2* expression being downregulated. Additionally, due to the complementary dynamic of these genes, *Gata2* might be downregulated in the cells where *Runx1* is expressed instead.

Both *RUNX1* and *GATA2* are commonly implemented markers in gene expression analysis for evaluating HSPCs and blood cell lineage emergence. The data presented indicates a complementary relationship between these genes, consistent with prior research. Findings have also shown that *GATA2* functions upstream of *RUNX1* as an early marker in the emergence of HSCs [43], which aligns with the expression mapping in the UMAPs presented in Figures 6g, h. The expression analysis of Exp EHB1 and EHB2 indicates this relationship. After the pre-treatment priming of the EBs, the expression at D0 and forward suggests an earlier upregulation of *GATA2*, with a later subsequent increase of *RUNX1* (Figures 6b, e).

## 3.5 Differential Impact on Caudal Transcription Factor Across Experimental Conditions

The expression of *HOXA9* in Exp IO displayed variability between CHIR conditions not seen in Exp EHB1 and EHB2, as illustrated in Figures 7a, b. In Exp IO, higher concentrations of CHIR promoted increased *HOXA9* expression, while the expression levels were more consistent across different CHIR conditions in Exp EHB1 and EHB2. When analyzing the pooled data, a pattern was observed where higher CHIR concentrations consistently enhanced *HOXA9* expression at each time point (Figure 7c). Despite this trend, the variability in *HOXA9* upregulation between Exp IO and the other experiments contributed to a high variance in the combined results, particularly for the 5 µM condition at D8 and D13.

The UMAP analyses in Figures 7d, e demonstrate the diverse expression patterns of *HOXA9* across stages of embryogenesis and in leukemia, respectively. As presented in Section 1.1.2, *HOXA9* serves as a marker for HSPCs and caudal transcription. This association is evident in Figure 7d, where *Hoxa9* expression is detected in various mesodermal lineages, including caudal mesoderm. In the leukemia dataset, *Hoxa9* expression is more uniformly distributed, observed in both mature and immature hematopoietic cells (Figures 7e and S7d, e).

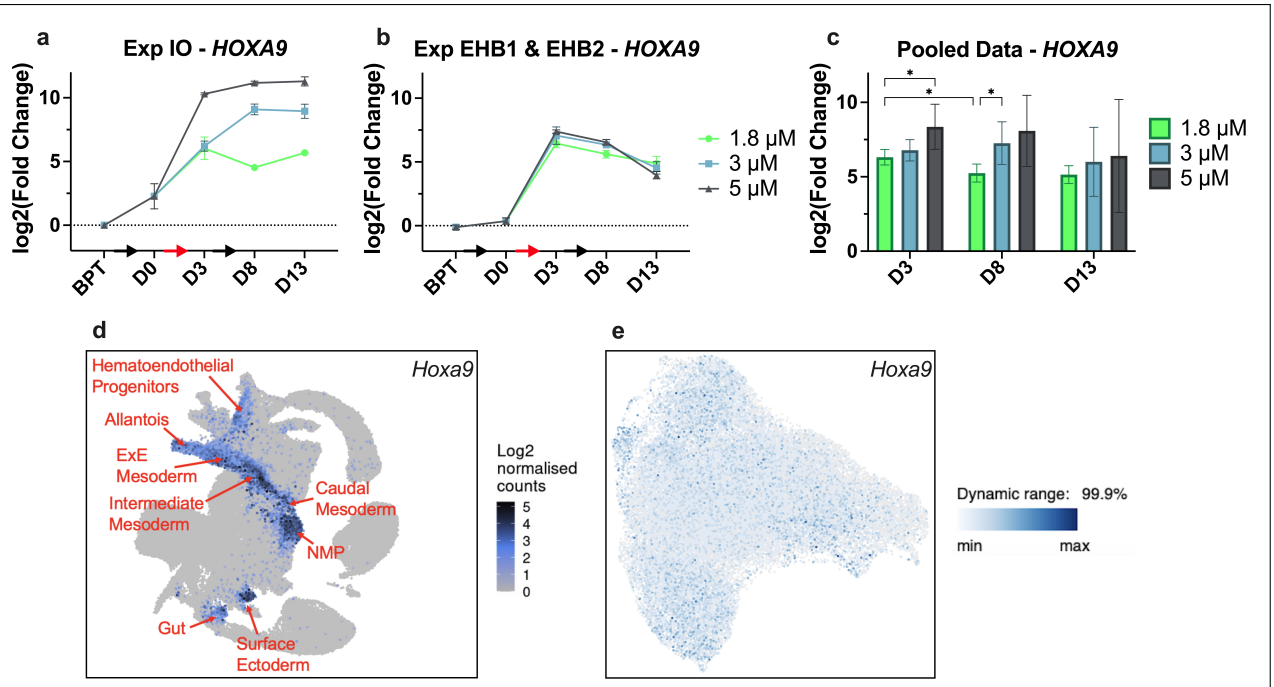


**Figure 7: Expression Profile of Caudal Transcription Marker**

**a,** *HOXA9* expression measured by RT-qPCR in Exp IO over time. BPT with n=1; remaining time points with n=2 for each. **b,** *HOXA9* expression in Exp EHB1 and EHB2 over time. D0 with n=2, and remaining time points with n=4 (2 independent experiments) for each. Black arrow: 3 µM CHIR. Red arrow: 1.8, 3 or 5 µM CHIR. **c,** Combined *HOXA9* expression from Exp IO, EHB1, and EHB2, with n=6 per time point. Plots display $\log2(2^{-\Delta\Delta Ct})$. Values represent mean ± SD. Asterisks indicate levels of statistical significance as determined by two-way ANOVA and Tukey multiple comparisons test (*p≤0.05). **d,** UMAP of single-cell RNAseq data of mouse gastrulation and early organogenesis from [58], colored by *Hoxa9* expression level. Embryonic cell populations with notable gene expression are outlined in red. ExE: Extra-embryonic, NMP: Neuromesodermal progenitors. See also Fig. S6. **e,** UMAP of single-cell RNAseq data of a mouse model of MLL-fusion leukemia from [57], colored by *Hoxa9* expression level. See also Fig. S7a.

These results suggest that CHIR promotes *HOXA9* expression, as demonstrated by the increase in expression following the addition of CHIR. Notably, higher concentrations of CHIR led to more pronounced upregulation, a result only observed in Exp IO data. This pattern may indicate that a higher CHIR concentration drives more significant differentiation towards caudal mesoderm lineages, an association illustrated in Figure 7d. However, such effects cannot be determined for Exp EHB1 and EHB2, where the influence of CHIR is less pronounced across different conditions.

## 3.6 HSPC Marker *SPINK2* Showed Almost Unchanged Expression

Figure 8a presents the expression of *SPINK2* in Exp IO. The gene was observed to increase slightly before being downregulated at later time points, particularly for the 5 µM condition. The UMAP plot of gastrulation and early organogenesis (Figure 8b) displayed *Spink2* to be

predominantly expressed in the epiblast lineage. No particular pattern of *Spink2* expression could be seen in the leukemia dataset, as plotted in Figure 8c.

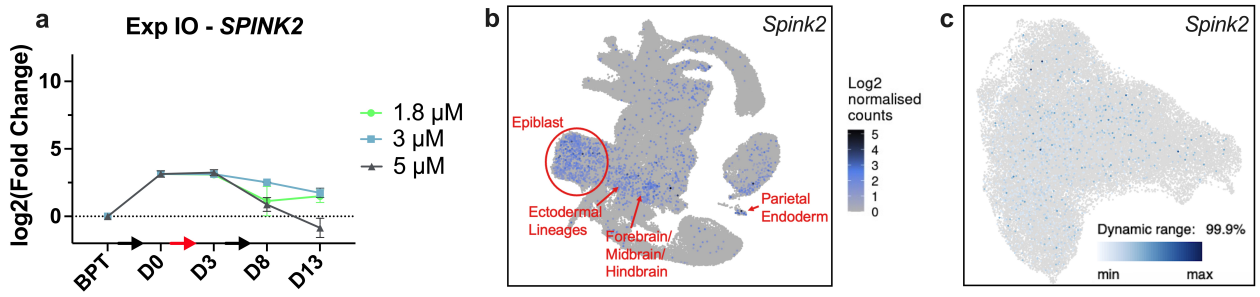


**Figure 8: Expression Profile of *SPINK2* as HSC Marker**

**a,** *SPINK2* expression measured by RT-qPCR in Exp IO over time. BPT with n=1; remaining time points with n=2 for each. Black arrow: 3 µM CHIR. Red arrow: 1.8, 3 or 5 µM CHIR. Plots display $\log2(2^{-\Delta\Delta Ct})$. Values represent mean ± SD. **b,** UMAP of single-cell RNAseq data of mouse gastrulation and early organogenesis from [58], colored by *Spink2* expression level. Embryonic cell populations with notable gene expression are outlined in red. See also Fig. S6. **c,** UMAP of single-cell RNAseq data of a mouse model of MLL-fusion leukemia from [57], colored by *Spink2* expression level. See also Fig. S7a.

These results suggest that *SPINK2* might be restricted to a smaller subset of HSCs, as indicated by the almost unchanged expression level of the gene across conditions. This is in line with previous findings which have shown *SPINK2*, a less established HSPC marker, to be a expressed in an HSC subset [46]. The association with the epiblast lineage might also point to *SPINK2* being a marker for pluripotency. An additional consideration is that as RT-qPCR is a bulk analysis method, this might not capture the expression of the gene in the HSCs within the EBs in an optimal way.

Given the limited variation in *SPINK2* expression in Exp IO, for the subsequent experiments (EHB1 and EHB2) it was decided to replace *SPINK2* with *CDH5* analysis. This was made in order to have a specific endothelial marker, to assess whether *SOX17* expression indicated the development of endothelium rather than that of endoderm.

## 3.7 Relative Expression of Genes Underlines Hypothesis about Lineage Emergence

Figures 9a-c display the relative expression of all markers utilized in the RT-qPCR analysis across all three experiments combined. The relative expression of *RSPO3* increased significantly until D3, followed by a slight stabilization and then a decrease. This trend was more evident for the two higher concentrations as illustrated by Figures 9b, c. Notably, the relative expression for *SOX17* and *CDH5* followed a similar trend, with a peak at D8 across all conditions. Similarly, *RUNX1* and *GATA2* exhibited consistent trends across all CHIR concentrations. *RUNX1* reached a higher relative expression at D0 compared to *GATA2*, but by D3, *GATA2* surpassed *RUNX1* in expression, with both genes peaking at D13. The relative expression of *HOXA9* was highest at D3, except in the 3 µM condition, and decreased at subsequent time points. *SPINK2*,

included in the Exp IO data analysis as an HSC marker, showed a relative expression diverging from the other genes. As *SPINK2* is expressed in epiblasts it might mark for pluripotent cells. The observed decrease in SPINK2 expression after D0 across all conditions may therefore indicate a loss of pluripotency.

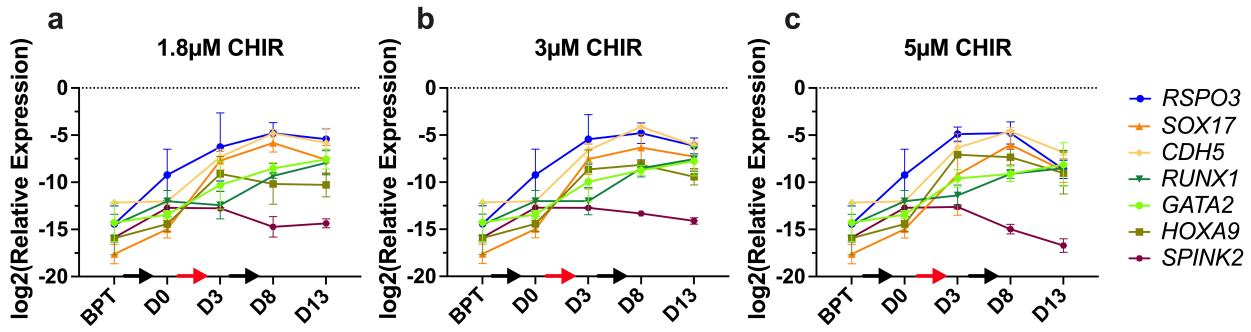


**Figure 9: Relative Expression of All Genes in Pooled Data**

**a, b, c,** RT-qPCR analysis showing pooled relative expression of gene markers *RSPO3, SOX17, RUNX1, GATA2, HOXA9, SPINK2,* and *CDH5* under different CHIR treatments in Exp IO, EHB1, and EHB2. Time points: BPT with n=5; D0 with n=4; and EBs at D3, D8, and D13 with n=6 for each. Exceptions are *SPINK2* with n=1 for BPT and n=2 for remaining time points, and *CDH5* with n=2 for D0, and n=4 for remaining time points. Panels **a, b,** and **c** correspond to conditions treated with 1.8µM, 3µM, and 5µM CHIR, respectively. Black arrow: 3 µM CHIR. Red arrow: 1.8, 3 or 5 µM CHIR. Values represent mean ± SD. Plots display $\log2(2^{-\Delta Ct})$ levels of pooled data from Exp IO, EHB1 and EHB2.

The expected lineage emergence hypothesized, as displayed in Figure 3f, appears to be validated by these results. The initial increase in expression of *RSPO3* after pre-treatment with CHIR precedes the other gene markers. This is followed by an emergence of endothelium lineage, indicated by increased expression of *SOX17* and *CDH5* by D3, across all conditions except the 5 µM condition. Subsequently, markers of HSPC rose, suggesting a transition from endothelial to hematopoietic lineage.

## 3.8 Endothelial and Blood Lineages Confirmed by Flow Cytometry

In order to verify the results produced from the RT-qPCR analysis, flow cytometry was done to evaluate the protein expression profiles of the cells. Via this analysis method, blood and endothelial cell populations were evaluated. Blood cells were gated using the pan-hematopoietic markers CD43 and CD45, while endothelial cells were marked by VE-cadherin (CDH5). In Figure 10a the gating strategy for these populations is presented using the 3 µM condition at D8 and D13 as reference. For the full gating strategy of the same reference sample, see Figure S8. The blood cells significantly increased, and the endothelial cells significantly decreased over time, displayed in Figures 10b and 10c as percentages of the live cell population, respectively. Notably, the 5 µM CHIR condition exhibited a greater magnitude of change in both cell populations between the two time points. This is also seen in the percentage of blood cells comparing the conditions at D13.

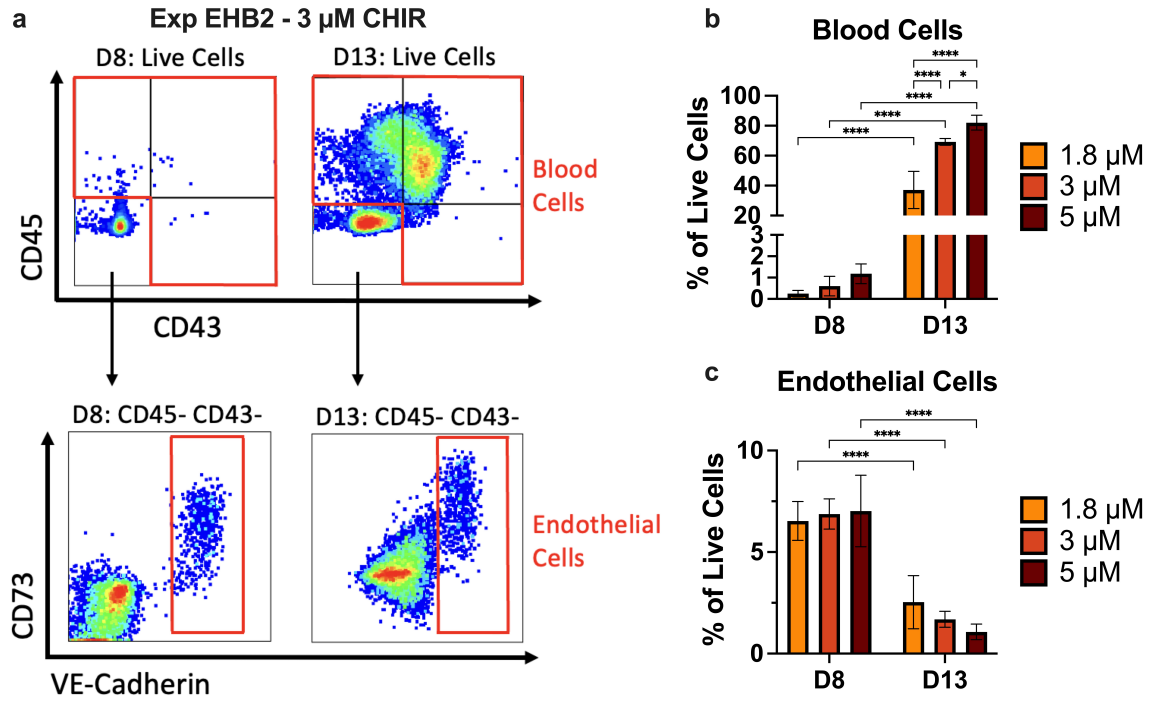


**Figure 10: Flow Cytometry Analysis of Endothelial and Blood Cell Populations**

**a,** Representative flow cytometry gating strategy for identifying blood (upper plots) and endothelial (lower plots) cell populations on D8 and D13 in the 3 µM condition (Exp EHB2). Blood cells are identified within the live cell population using CD45 and CD43 markers. Endothelial cells are gated within the CD45- CD43- population, further identified with VE-Cadherin as marker. Target populations highlighted in red. **b, c,** Percentage of blood cells (**b**) and endothelial cells (**c**) within the live cell population at D8 and D13 across the three conditions, with n=4 for both time points. Values represent mean ± SD. Asterisks indicate levels of statistical significance as determined by two-way ANOVA and Tukey multiple comparisons test (*p≤0.05, ****p≤0.0001).

These results confirm the emergence of endothelial and blood cell populations as indicated by the RT-qPCR analysis. Specifically, the blood cell population increased over time, a trend that was consistently observed in both the gene expression and protein expression assays. In contrast, the endothelial cell population showed a peak at day 8 in the RT-qPCR data, confirmed by flow cytometry, with a decline by day 13. Interestingly, the difference in cell populations between conditions was more pronounced in the flow cytometry data, as opposed to the lineage emergence suggested by the gene expression measurements.

# 4

# Discussion

This section starts off with a main discussion in regard to the results presented in this thesis. Thereafter, some limitations to the project are addressed. Finally, different wider implications are explored, relating to possible future work that can be implemented to continue the research of the project.

## 4.1 Main Discussion in Relation to the Project Aims

The discussion is centered around the aims of the project and the related questions, as presented in Section 1.2.1. For each question, the results will be discussed in an overarching manner.

**Aim 1: Confirm Wnt Pathway Activation and Optimize Modulation Conditions**

- Can Wnt pathway activation be effectively confirmed in the generated embryoid bodies, and how do different concentrations of CHIR influence this activation?

In all experiments, *RSPO3* was upregulated after the addition of CHIR, which verifies an activation of the Wnt signaling pathway in the generated EBs. Hence leading to the differentiation of cells and lineage emergence as confirmed by the remaining gene expression results. Endothelial lineage emerged, as marked by *SOX17* and *CDH5*, as well did the remaining HSPC markers indicate that subsequent hematopoiesis arose within the EBs.

Exp IO data displayed a notable effect between conditions. The expression of *RSPO3* was however not varied between conditions at all time points post-CHIR addition. The most evident difference was seen directly after D0, showing an effect to CHIR, which was then less pronounced when the same concentration of treatment was added at D3. This indicates that the different levels of CHIR had an effect on Wnt signaling in this experiment, as the *RSPO3* expression became more consistent when 3 µM CHIR was added to all EBs.

Looking more specifically at the effect of CHIR, and therefore possibly Wnt signaling, the upregulation of *HOXA9* and *SOX17* in Exp IO showed an inverted relationship between conditions. A higher CHIR concentration significantly upregulated *HOXA9* and downregulated *SOX17*. These results suggest that higher concentrations of CHIR enhance differentiation towards mesoderm lineages such as HSPCs, while lower concentrations promote differentiation towards endothelial cells.

In conclusion, the Wnt signaling pathway was activated within the EBs in all three experiments. The variance between conditions was less pronounced in Exp EHB1 and EHB2 than in Exp IO, nevertheless, Wnt signaling was activated and subsequent lineages emerged as promoted by the signaling pathway.

**Aim 2: Evaluate the Specificity and Efficiency of the Established Protocol**

- Does the established protocol consistently give rise to the desired cell types, specifically endothelial cells and hematopoietic stem and progenitor cells, within the generated embryoid bodies?

The evaluated genes were upregulated in the generated EBs, reflecting successful lineage differentiation. Notably, the endothelial lineage emerged and peaked by D8 in all experiments, validated by the expression of *SOX17* and *CDH5*. This reflects the specificity of the established protocol in consistently giving rise to the desired endothelium lineage. Significant upregulation was observed across all experiments for the HSPC genes as well. However, there was a notable variability in the expression levels of some HSPC markers between experiments, particularly *HOXA9, RUNX1*, and *GATA2*.

*HOXA9* gene expression indicated differentiation towards hematopoietic lineages and a possible A-P axis definition [28]. In Exp IO the effect of CHIR on *HOXA9* expression was more pronounced, while in Exp EHB1 and EHB2, the conditions had less effect. In the later experiments *HOXA9* expression declined at later time points. This might have had an impact on elongation, as suggested by the uniformly spherical morphology of the EBs (Figures 3c and S5).

Interestingly, as *HOXA9* decreased in Exp EHB1 and EHB2, *RUNX1* and *GATA2* levels increased. This increase was also more evident in these experiments in contrast to Exp IO, where the upregulation of *RUNX1* and *GATA2* was less pronounced, while *HOXA9* was more upregulated at later stages. This shift might suggest that, in Exp EHB1 and EHB2, genes involved in blood cell lineage development were more prominently upregulated than those involved in caudal transcription.

As verified by protein expression assays through flow cytometry analysis, both endothelial and blood cell lineages were successfully derived within the EBs. Consistent with expectations, endothelial cells predominated at earlier stages, exhibiting a higher population across all conditions by D8 compared to D13. In line with this, the population of blood cell lineage increased steadily, reaching its peak at the final time point on D13. Notably, as discussed in Results Section 3.8), the differences in blood cell populations between conditions at D13 were more pronounced than those suggested by RT-qPCR results. This discrepancy likely stems from the function of RT-qPCR as a bulk analysis technique, which can dilute individual gene expression variations. This sets the stage for why flow cytometry is a valuable tool to incorporate, as it provides a single-cell analysis perspective to accurately profile dynamic cellular processes within the EBs.

Ultimately, while the specificity of the protocol in inducing the same HSPC lineages showed some variation between Exp IO, and the Exp EHB1 and EHB2, all experiments demonstrated the emergence of HSPCs over time. Notably, the low variance observed in gene expression in the Exp EHB1 and EHB2 dataset, suggests a specificity of the protocol to give rise to desired hematopoietic lineages in these later experiments.

**Aim 3: Assess the Reproducibility of the Differentiation Protocol**

- How reproducible are the differentiation outcomes when utilizing the same established protocol with consistent CHIR concentrations across multiple experiments?

The gene expression profiles from Exp IO displayed notable variability between CHIR conditions

when compared to Exp EHB1 and EHB2, indicating differences in experimental outcomes. In Exp IO, the expression of the genes was more affected by the experimental conditions compared to Exp EHB1 and EHB2.

The variation between experiments could be attributed to several factors, such as the modifications in the protocol regarding the CHIR addition method at D0. As described in Section 2.1.3, varying volumes from the same CHIR stock (15 mM) were used in Exp IO, which seems to have introduced more pronounced variability in the treatment effects. In contrast, in Exp EHB1 and EHB2, a standardized volume addition of CHIR stocks of different initial concentrations was employed.

Other critical factors to consider are the analytical methods employed and the variability introduced by different operators conducting the same protocol. As mentioned, while bulk analysis techniques such as RT-qPCR provide a cumulative view of gene expression across a population of cells, they can mask individual cell behaviors and subtle dynamics within the cell population. Regarding the latter factor, Exp IO was carried out by a previous student, while I conducted Exp EHB1 and EHB2. Subtle differences in protocol execution might have been introduced due to this, contributing to the observed variations in gene expression.

All things considered, the established protocol utilized in the lab group as of now seems to be specific and efficient in the differentiation of HSPCs, producing similar results indicative of a reproducible experimental protocol.

## 4.2 Limitations of the Project

Throughout this project, various analytical methods were employed to evaluate the protocols and the generated EBs. Although the analyses were comprehensive, certain limitations were imposed due to time constraints.

For the relative quantification of RT-qPCR data, a simplified version of the Pfaffl method was used (the Livak method). This approach presumes optimal primer efficiency at 100%, yet the accepted range of primer efficiencies is 90-105% [56]. After primer optimization, some of the primers used in this project had efficiencies outside of the accepted range. By evaluating the dissociation curves for each primer after each RT-qPCR analysis, specific amplification was still carefully controlled and determined. Nevertheless, it would have been interesting to employ the more precise Pfaffl method parallel to the simplified Livak method as a way to evaluate each method and the results generated.

Additionally, the marker for *CDH5* expression was introduced only in Exp EHB1 and EHB2 to confirm the emergence of endothelial cells. It would have been beneficial to include RT-qPCR analysis using this marker in the Exp IO dataset as well, to verify the endothelial lineage. Unfortunately, time constraints limited the scope of these investigations as well.

## 4.3 Wider Implications and Future Directions

Throughout this thesis, the modulation of Wnt pathway activity using CHIR, and its effect on cell differentiation, have been central. The roles of different genes in evaluating lineage emergence have also been crucial. These aspects are particularly significant because they directly contribute to our understanding of hematopoietic processes, which are pivotal in the strive

to understand hematopoietic disruptions. As highlighted in Section 1.1, hiPSC-derived EBs serve as a critical model system for studying leukemia. Therefore, optimizing these systems is an essential step forward in advancing our understanding and developing interventions for hematopoietic disorders.

Firstly, studying Wnt signaling further would give a deeper insight into how it promotes blood development. This could for instance be done by optimizing the addition of CHIR in regard to the timing and concentration of the CHIR addition in the protocol. Besides this, instead of studying the *RSPO3* expression, it would be interesting to look into other Wnt signaling markers that might prove more optimal. Another perspective would be to investigate the cell membrane receptors of Wnt proteins, namely the different Frizzled proteins or the lipoprotein receptor-related protein (LRP) [63], as a way to evaluate the signaling activation. Similarly, studying the secreted Wnt proteins would be a possible way forward.

Additionally, some of the genes of interest to the project have been shown to play essential roles in leukemia. For instance, findings have shown that *RSPO3* and *HOXA9* are tightly connected to certain leukemia types [64]. As *RSPO3* enhances Wnt/$\beta$-catenin signaling, it is critical for leukemia stem cell self-renewal, particularly in AML with MLL translocations which are notably prevalent in infants. *HOXA9* is often overexpressed in AML, acting as a poor prognostic indicator and contributing to leukemic transformation. Incorporating these factors in future work could advance targeted strategies for treating leukemia.

Continuing on the perspective of leukemia and other possible ways to further study this hematopoietic disruption, it would be interesting to generate EBs utilizing hiPSCs carrying childhood leukemia-associated mutations. By doing so, the EBs could serve a dual purpose. Besides being used to study hematopoiesis and optimize HSPC production, the EBs could also provide crucial insights into the initiation and progression of leukemia

Both gene expression analysis by RT-qPCR and protein expression assay via flow cytometry were employed in this thesis, each offering distinct advantages. While flow cytometry provides valuable single-cell insights into specific cell populations, it is limited to the specific antibodies available. To gain a deeper understanding of the cellular dynamics within the EBs, employing scRNA-seq could be highly beneficial. This could provide a richer and more nuanced understanding of cellular processes and differentiation patterns in the context of hematopoietic development in the EBs.

Lastly, utilizing the EBs in the development and production of HSCs with long-term engraftment potential is an important future direction to take this research. As noted in the literature, the expansion of HSCs remains a critical challenge due to their limited growth potential and tendency to differentiate in culture [53]. There is potential to overcome these hurdles by implementing the expansion of HSCs utilizing EBs. This approach could in the future make use of the controlled differentiation environment of EBs to maintain and even expand HSCs under serum-free and well-defined conditions.

In conclusion, this project is tied to a range of different wider implications, and the various approaches to develop the research further are many. Whether it is looking further into the genes that have been utilized to evaluate the generated EBs, the Wnt signaling pathway as a key hematopoiesis enhancer, or the actual EBs serving as model systems for blood development, various directions can be taken to deepen our understanding of the processes involved in embryonic development.

# Bibliography


[1] A. Chennamadhavuni, V. Lyengar, S. Mukkamalla, and A. Shimanovsky, *Leukemia*. StatPearls Publishing, 2023. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/32809325/.

[2] J. Huang, S. Chan, C. Ngai, *et al.*, "Disease Burden, Risk Factors, and Trends of Leukaemia: A Global Analysis," *Frontiers in oncology*, vol. 12, no. 904292, 2022. DOI: 10.3389/fonc.2022.904292.

[3] D. Shanmugavadivel, J.-F. Liu, A. Ball-Gamble, *et al.*, "The Childhood Cancer Diagnosis (CCD) Study: a UK observational study to describe referral pathways and quantify diagnostic intervals in children and young people with cancer," *BMJ Open*, vol. 12, no. 2, 2022, ISSN: 2044-6055. DOI: 10.1136/bmjopen-2021-058744.

[4] S. M. Namayandeh, Z. Khazaei, M. Lari Najafi, E. Goodarzi, and A. Moslem, "GLOBAL Leukemia in Children 0-14 Statistics 2018, Incidence and Mortality and Human Development Index (HDI): GLOBOCAN Sources and Methods," *Asian Pacific Journal of Cancer Prevention*, vol. 21, no. 5, pp. 1487–1494, 2020, ISSN: 1513-7368. DOI: 10.31557/APJCP.2020.21.5.1487.

[5] "Childhood cancer." (2021), [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer-in-children#:~:text=In%20December%202021%2C%20WHO%20and,dispensing%20medicines%20according%20to%20best.

[6] A. Cazzola, G. Cazzaniga, A. Biondi, R. Meneveri, S. Brunelli, and E. Azzoni, "Prenatal Origin of Pediatric Leukemia: Lessons From Hematopoietic Development," *Frontiers in Cell and Developmental Biology*, vol. 8, 2021, ISSN: 2296-634X. DOI: 10.3389/fcell.2020.618164.

[7] G. Marshall, D. Carter, and B. e. a. Cheung, "The prenatal origins of cancer," *Nat Rev Cancer*, vol. 14, no. 4, pp. 277–289, 2014. DOI: 10.1038/nrc3679.

[8] D. Hein, A. Borkhardt, and U. Fischer, "Insights into the prenatal origin of childhood acute lymphoblastic leukemia," *Cancer Metastasis Rev*, vol. 39, no. 1, pp. 161–171, 2020. DOI: 10.1007/s10555-019-09841-1.

[9] C. Meyer, T. Burmeister, D. Gröger, *et al.*, "The MLL recombinome of acute leukemias in 2017," *Leukemia*, vol. 32, Jul. 2017. DOI: 10.1038/leu.2017.213.

[10] C. Gerri, S. Menchero, S. K. Mahadevaiah, J. M. Turner, and K. K. Niakan, "Human embryogenesis: A comparative perspective," *Annual Review of Cell and Developmental Biology*, vol. 36, no. Volume 36, 2020, pp. 411–440, 2020, ISSN: 1530-8995. DOI: 10.1146/annurev-cellbio-022020-024900.

[11] M. N. Shahbazi, "Mechanisms of human embryo development: from cell fate to tissue shape and back," *Development*, vol. 147, no. 14, dev190629, Jul. 2020, ISSN: 0950-1991. DOI: 10.1242/dev.190629. eprint: https://journals.biologists.com/dev/article-pdf/147/14/dev190629/1967449/dev190629.pdf.

[12] P. Blakeley, N. M. E. Fogarty, I. del Valle, *et al.*, "Defining the three cell lineages of the human blastocyst by single-cell RNA-seq," *Development*, vol. 142, no. 18, pp. 3151–3165, Sep. 2015, ISSN: 0950-1991. DOI: 10.1242/dev.123547. eprint: https://journals.biologists.com/dev/article-pdf/142/18/3151/1838771/dev123547.pdf.

[13] S. Ghimire, V. Mantziou, N. Moris, and A. Martinez Arias, "Human gastrulation: The embryo and its models," *Developmental Biology*, vol. 474, pp. 100–108, 2021, Synthetic Embryology, ISSN: 0012-1606. DOI: 10.1016/j.ydbio.2021.01.006.

[14] A. T. Hertig, J. Rock, and E. C. Adams, "A description of 34 human ova within the first 17 days of development," *American Journal of Anatomy*, vol. 98, no. 3, pp. 435–493, 1956. DOI: 10.1002/aja.1000980306. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aja.1000980306.

[15] W.-H. Tan, E. C. Gilmore, and H. N. Baris, "Chapter 15 - Human Developmental Genetics," in *Emery and Rimoin's Principles and Practice of Medical Genetics (Sixth Edition)*, D. Rimoin, R. Pyeritz, and B. Korf, Eds., Sixth Edition, Oxford: Academic Press, 2013, pp. 1–63, ISBN: 978-0-12-383834-6. DOI: 10.1016/B978-0-12-383834-6.00018-5.

[16] G. M. Anand, H. C. Megale, S. H. Murphy, *et al.*, "Controlling organoid symmetry breaking uncovers an excitable system underlying human axial elongation," *Cell*, vol. 186, no. 3, 497–512.e23, 2023, ISSN: 0092-8674. DOI: 10.1016/j.cell.2022.12.043.

[17] B. T. MacDonald, K. Tamai, and X. He, "Wnt/$\beta$-Catenin Signaling: Components, Mechanisms, and Diseases," *Developmental Cell*, vol. 17, no. 1, pp. 9–26, 2009, ISSN: 1534-5807. DOI: 10.1016/j.devcel.2009.06.016.

[18] P. Andre, H. Song, W. Kim, A. Kispert, and Y. Yang, "Wnt5a and Wnt11 regulate mammalian anterior-posterior axis elongation," *Development*, vol. 142, no. 8, pp. 1516–1527, Apr. 2015, ISSN: 0950-1991. DOI: 10.1242/dev.119065. eprint: https://journals.biologists.com/dev/article-pdf/142/8/1516/1841641/dev119065.pdf.

[19] H. Acloque, M. S. Adams, K. Fishwick, M. Bronner-Fraser, and M. A. Nieto, "Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease," *The Journal of Clinical Investigation*, vol. 119, no. 6, pp. 1438–1449, Jun. 2009. DOI: 10.1172/JCI38019. [Online]. Available: https://www.jci.org/articles/view/38019.

[20] H. Goto, S. C. Kimmey, R. H. Row, D. Q. Matus, and B. L. Martin, "FGF and canonical Wnt signaling cooperate to induce paraxial mesoderm from tailbud neuromesodermal progenitors through regulation of a two-step epithelial to mesenchymal transition," *Development*, vol. 144, no. 8, pp. 1412–1424, Apr. 2017, ISSN: 0950-1991. DOI: 10.1242/dev.143578. eprint: https://journals.biologists.com/dev/article-pdf/144/8/1412/1850991/dev143578.pdf.

[21] T. J. Cunningham, S. Kumar, T. P. Yamaguchi, and G. Duester, "Wnt8a and Wnt3a cooperate in the axial stem cell niche to promote mammalian body axis extension," *Developmental Dynamics*, vol. 244, no. 6, pp. 797–807, 2015. DOI: 10.1002/dvdy.24275. eprint: https://anatomypubs.onlinelibrary.wiley.com/doi/pdf/10.1002/dvdy.24275.

[22] J. Richter, D. Traver, and K. Willert, "The role of Wnt signaling in hematopoietic stem cell development," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 52, no. 4, pp. 414–424, 2017. DOI: 10.1080/10409238.2017.1325828.

[23] Y. Komiya and R. Habas, "Wnt signal transduction pathways," *Organogenesis*, vol. 4, no. 2, pp. 68–75, 2008. DOI: 10.4161/org.4.2.5851. eprint: https://doi.org/10.4161/org.4.2.5851.

[24] A. Ramakrishnan and K. M. Cadigan, "Wnt target genes and where to find them," *F1000Research*, vol. 6, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:3635409.

[25] M. Zhang, M. Haughey, N.-Y. Wang, *et al.*, "Targeting the Wnt signaling pathway through R-spondin 3 identifies an anti-fibrosis treatment strategy for multiple organs," *PLOS ONE*, vol. 15, no. 3, pp. 1–21, Mar. 2020. DOI: 10.1371/journal.pone.0229445.

[26] B. Ohkawara, A. Glinka, and C. Niehrs, "Rspo3 Binds Syndecan 4 and Induces Wnt/PCP Signaling via Clathrin-Mediated Endocytosis to Promote Morphogenesis," *Developmental Cell*, vol. 20, no. 3, pp. 303–314, 2011, ISSN: 1534-5807. DOI: 10.1016/j.devcel.2011.01.006.

[27] J. Pearson, D. Lemons, and W. Mcginnis, "Modulating Hox gene functions during animal body patterning," *Nature reviews. Genetics*, vol. 6, pp. 893–904, Jan. 2006. DOI: 10.1038/nrg1726.

[28] C. Li, S. Wright, X. Zhao, S. Mryncza, and B. Xu, "Systematically Interrogating Downstream Regulation Mechanisms of HOXA9 in MLL-r Leukemia Via Functional Noncoding CRISPR Screens," *Blood*, vol. 140, no. Supplement 1, pp. 1269–1270, Nov. 2022, ISSN: 0006-4971. DOI: 10.1182/blood-2022-159186.

[29] M. Epstein, G. Pillemer, R. Yelin, J. K. Yisraeli, and A. Fainsod, "Patterning of the embryo along the anterior-posterior axis: the role of the caudal genes," *Development*, vol. 124, no. 19, pp. 3805–3814, Oct. 1997, ISSN: 0950-1991. DOI: 10.1242/dev.124.19.3805. eprint: https://journals.biologists.com/dev/article-pdf/124/19/3805/3127047/develop_124_19_3805.pdf.

[30] N. Di-Poï, U. Koch, F. Radtke, and D. Duboule, "Additive and global functions of HoxA cluster genes in mesoderm derivatives," *Developmental Biology*, vol. 341, no. 2, pp. 488–498, 2010, ISSN: 0012-1606. DOI: 10.1016/j.ydbio.2010.03.006.

[31] M. Lambert, M. Alioui, S. Jambon, S. Depauw, I. V. Seuningen, and M.-H. David-Cordonnier, "Direct and Indirect Targeting of HOXA9 Transcription Factor in Acute Myeloid Leukemia," *Cancers*, vol. 11, no. 6, 2019, ISSN: 2072-6694. DOI: 10.3390/cancers11060837.

[32] A. Rowan-Hull, "Human organogenesis," in *Textbook of Clinical Embryology*, K. Coward and D. Wells, Eds. Cambridge University Press, 2013, pp. 118–132.

[33] G. Canu and C. Ruhrberg, "First blood: the endothelial origins of hematopoietic progenitors," *Angiogenesis*, vol. 24, no. 2, pp. 199–211, 2021. DOI: 10.1007/s10456-021-09783-9.

[34] E. Julien, R. El Omar, and M. Tavian, "Origin of the hematopoietic system in the human embryo," *FEBS Letters*, vol. 590, no. 22, pp. 3987–4001, 2016. DOI: 10.1002/1873-3468.12389. eprint: https://febs.onlinelibrary.wiley.com/doi/pdf/10.1002/1873-3468.12389.

[35] E. Dzierzak and A. Bigas, "Blood Development: Hematopoietic Stem Cell Dependence and Independence," *Cell Stem Cell*, vol. 22, no. 5, pp. 639–651, 2018, ISSN: 1934-5909. DOI: 10.1016/j.stem.2018.04.015.

[36] B. Hadland and M. Yoshimoto, "Many layers of embryonic hematopoiesis: new insights into B-cell ontogeny and the origin of hematopoietic stem cells," *Experimental Hematology*, vol. 60, pp. 1–9, 2018, ISSN: 0301-472X. DOI: 10.1016/j.exphem.2017.12.008.

[37] J. M. Frame, K. E. McGrath, and J. Palis, "Erythro-myeloid progenitors: "Definitive" hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells," *Blood Cells, Molecules, and Diseases*, vol. 51, no. 4, pp. 220–225, 2013, Embryonic and Fetal Hematopoiesis, ISSN: 1079-9796. DOI: 10.1016/j.bcmd.2013.09.006.

[38] K. Kissa and P. Herbomel, "Blood stem cells emerge from aortic endothelium by a novel type of cell transition," *Nature*, vol. 464, no. 7285, 2010. DOI: 10.1038/nature08761.

[39] C. Guibentif, R. E. Rönn, C. Böiers, *et al.*, "Single-Cell Analysis Identifies Distinct Stages of Human Endothelial-to-Hematopoietic Transition," *Cell Reports*, vol. 19, no. 1, pp. 10–19, 2017, ISSN: 2211-1247. DOI: 10.1016/j.celrep.2017.03.023.

[40] K. Ottersbach, "Endothelial-to-haematopoietic transition: an update on the process of making blood," *Biochemical Society Transactions*, vol. 47, no. 2, pp. 591–601, Mar. 2019, ISSN: 0300-5127. DOI: 10.1042/BST20180320. eprint: https://portlandpress.com/biochemsoctrans/article-pdf/47/2/591/850907/bst-2018-0320c.pdf.

[41] R. L. Clarke, A. D. Yzaguirre, Y. Yashiro-Ohtani, *et al.*, "The expression of Sox17 identifies and regulates haemogenic endothelium.," *Nature Cell Biology*, vol. 15, no. 5, pp. 502–510, 2013, ISSN: 14657392. [Online]. Available: https://search.ebscohost.com/login.aspx?direct=true&db=asx&AN=87383334&site=eds-live&scope=site&authtype=guest&custid=s3911979&groupid=main&profile=eds.

[42] A. D. Yzaguirre, M. F. T. R. de Bruijn, and N. A. Speck, "The role of runx1 in embryonic blood cell formation," in *RUNX Proteins in Development and Cancer*, Y. Groner, Y. Ito, P. Liu, J. C. Neil, N. A. Speck, and A. van Wijnen, Eds. Singapore: Springer Singapore, 2017, pp. 47–64, ISBN: 978-981-10-3233-2. DOI: 10.1007/978-981-10-3233-2_4.

[43] E. Bresciani, B. Carrington, K. Yu, *et al.*, "Redundant mechanisms driven independently by RUNX1 and GATA2 for hematopoietic development," *Blood Advances*, vol. 5, no. 23, pp. 4949–4962, Nov. 2021, ISSN: 2473-9529. DOI: 10.1182/bloodadvances.2020003969. eprint: https://ashpublications.org/bloodadvances/article-pdf/5/23/4949/1847600/advancesadv2020003969.pdf.

[44] K. R. Calvo and D. D. Hickstein, "The spectrum of GATA2 deficiency syndrome," *Blood*, vol. 141, no. 13, pp. 1524–1532, 2023, ISSN: 0006-4971. DOI: 10.1182/blood.2022017764.

[45] F. L. Bos, J. S. Hawkins, and A. C. Zovein, "Single-cell resolution of morphological changes in hemogenic endothelium," *Development*, vol. 142, no. 15, pp. 2719–2724, Aug. 2015, ISSN: 0950-1991. DOI: 10.1242/dev.121350. eprint: https://journals.biologists.com/dev/article-pdf/142/15/2719/1838040/dev121350.pdf.

[46] V. Calvanese, S. Capellera-Garcia, F. Ma, *et al.*, "Mapping human haematopoietic stem cells from haemogenic endothelium to birth," *Nature*, vol. 604, pp. 1–7, Apr. 2022. DOI: 10.1038/s41586-022-04571-x.

[47] K. Takahashi and S. Yamanaka, "Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors," *Cell*, vol. 126, no. 4, pp. 663–676, 2006, ISSN: 0092-8674. DOI: 10.1016/j.cell.2006.07.024.

[48] K. Zeevaert, M. H. Elsafi Mabrouk, W. Wagner, and R. Goetzke, "Cell Mechanics in Embryoid Bodies," *Cells*, vol. 9, no. 10, 2020, ISSN: 2073-4409. DOI: 10.3390/cells9102270.

[49] G. Pettinato, X. Wen, and N. Zhang, "Engineering Strategies for the Formation of Embryoid Bodies from Human Pluripotent Stem Cells," *Stem Cells and Development*, vol. 24, no. 14, pp. 1595–1609, 2015, PMID: 25900308. DOI: 10.1089/scd.2014.0427.

[50] N. Moris, K. Anlas, J. Schröder, *et al.*, "Generating Human Gastruloids from Human Embryonic Stem Cells," *Protocol Exchange*, Jun. 2020. DOI: 10.21203/rs.3.pex-812/v1.

[51] N. Moris, K. Anlas, S. C. van den Brink, *et al.*, "An in vitro model of early anteroposterior organization during human development," *Nature*, vol. 582, no. 7812, pp. 410–415, Jun. 2020, ISSN: 0028-0836. DOI: 10.1038/s41586-020-2383-9.

[52] D. Vestweber, "Ve-cadherin," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 28, no. 2, pp. 223–232, 2008. DOI: 10.1161/ATVBAHA.107.158014. eprint: https://www.ahajournals.org/doi/pdf/10.1161/ATVBAHA.107.158014.

[53] S. Bastani, F. Staal, and K. Canté-Barrett, "The quest for the holy grail: overcoming challenges in expanding human hematopoietic stem cells for clinical use," *Stem cell investigation*, vol. 10, p. 15, Jul. 2023. DOI: 10.21037/sci-2023-016.

[54] A. Ditadi and C. M. Sturgeon, "Directed differentiation of definitive hemogenic endothelium and hematopoietic progenitors from human pluripotent stem cells," *Methods*, vol. 101, pp. 65–72, 2016, Human Pluripotent Stem Cells, ISSN: 1046-2023. DOI: 10.1016/j.ymeth.2015.10.001.

[55] L. Oburoglu, E. Mansell, I. Canals, *et al.*, "Pyruvate metabolism guides definitive lineage specification during hematopoietic emergence," *EMBO reports*, vol. 23, Dec. 2021. DOI: 10.15252/embr.202154384.

[56] Bio-Rad Laboratories, Inc., *RT-qPCR Applications Guide*, Accessed online, Bio-Rad Laboratories, Inc., 2006. [Online]. Available: https://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_5279.pdf.

[57] D. Sjövall, S. Ghosh, J. Hansson, C. Guibentif, and P. Jaako, "Defective ribosome assembly impairs leukemia stem cell function in a murine model of acute myeloid leukemia," *bioRxiv*, 2023. DOI: 10.1101/2023.01.22.525120. eprint: https://www.biorxiv.org/content/early/2023/01/23/2023.01.22.525120.full.pdf.

[58] B. Pijuan-Sala, J. Griffiths, C. Guibentif, *et al.*, "A single-cell molecular map of mouse gastrulation and early organogenesis," *Nature*, Feb. 2019. DOI: 10.1038/s41586-019-0933-9.

[59] R. Nagasaka, M. Matsumoto, M. Okada, *et al.*, "Visualization of morphological categories of colonies for monitoring of effect on induced pluripotent stem cell culture status," *Regenerative Therapy*, vol. 6, pp. 41–51, 2017, ISSN: 2352-3204. DOI: 10.1016/j.reth.2016.12.003.

[60] Y. Wang, A. V. Krivtsov, A. U. Sinha, *et al.*, "The Wnt/$\beta$-Catenin Pathway Is Required for the Development of Leukemia Stem Cells in AML," *Science*, vol. 327, no. 5973, pp. 1650–1653, 2010. DOI: 10.1126/science.1186624. eprint: https://www.science.org/doi/pdf/10.1126/science.1186624.

[61] D. Láinez-González, A. B. Alonso-Aguado, and J. M. Alonso-Dominguez, "Understanding the Wnt Signaling Pathway in Acute Myeloid Leukemia Stem Cells: A Feasible Key against Relapses," *Biology*, vol. 12, no. 5, 2023, ISSN: 2079-7737. DOI: 10.3390/biology12050683.

[62] E. Oberlin, M. Fleury, D. Clay, *et al.*, "VE-cadherin expression allows identification of a new class of hematopoietic stem cells within human embryonic liver," *Blood*, vol. 116, no. 22, pp. 4444–4455, 2010, ISSN: 0006-4971. DOI: 10.1182/blood-2010-03-272625.

[63] B. Macdonald and X. He, "Frizzled and LRP5/6 Receptors for Wnt/ -Catenin Signaling," *Cold Spring Harbor perspectives in biology*, vol. 4, Dec. 2012. DOI: 10.1101/cshperspect.a007880.

[64] B. Salik, H. Yi, N. Hassan, *et al.*, "Targeting RSPO3-LGR4 Signaling for Leukemia Stem Cell Eradication in Acute Myeloid Leukemia," *Cancer Cell*, vol. 38, no. 2, 263–278.e6, 2020, ISSN: 1535-6108. DOI: 10.1016/j.ccell.2020.05.014.

# A

# Appendix I

**Table A.1:** Table displaying components added to each SP34 Medium

| Medium | Component | Final concentration |
|---|---|---|
| SP34-d3 | Glutamax | 1x |
| | Ascorbic acid | 50 µg/mL |
| | Transferrin (30 mg/mL) | 150 µg/mL |
| | BME 1000X | 55 µM |
| | VEGF | 15 ng/mL |
| | bFGF | 5 ng/mL |
| | CHIR | 3 µM |
| | Activin A | 1 ng/mL |
| SP34-d4 | Glutamax | 1x |
| | Ascorbic acid | 50 µg/mL |
| | Transferrin (30 mg/mL) | 150 µg/mL |
| | BME 1000X | 55 µM |
| | VEGF | 15 ng/mL |
| | bFGF | 5 ng/mL |
| SP34-d6 | Glutamax | 1x |
| | Ascorbic acid | 50 µg/mL |
| | Transferrin (30 mg/mL) | 150 µg/mL |
| | BME 1000X | 55 µM |
| | VEGF | 15 ng/mL |
| | bFGF | 5 ng/mL |
| | SCF | 200 ng/mL |
| | EPO | 4 IU/mL |
| | IL6 | 20 ng/mL |
| | IL11 | 10 ng/mL |
| | IGF1 | 50 ng/mL |
| SP34-d8 | Glutamax | 1x |
| | Ascorbic acid | 50 µg/mL |
| | Transferrin (30 mg/mL) | 150 µg/mL |
| | BME 1000X | 55 µM |
| | VEGF | 5 ng/mL |
| | bFGF | 5 ng/mL |
| | SCF | 100 ng/mL |
| | EPO | 2 IU/mL |
| | IL6 | 10 ng/mL |

Continued on next page

Table A.1 – Continued from previous page

| Medium | Component | Final Concentration |
|---|---|---|
| | IL11 | 5 ng/mL |
| | IGF1 | 25 ng/mL |
| | TPO | 30 ng/mL |
| | Flt3L | 10 ng/mL |
| | IL3 | 30 ng/mL |

**Table A.2:** Table displaying details regarding forward and reverse primers for each gene evaluated in the RT-qPCR analysis. This includes primer sequence, melting point, GC content, length of each primer, product size, and efficiency. ND: Not determined.

| Gene | Primer | Sequence | Melting Point [°C] | GC% | Length | Product Length | Efficiency (%) |
|---|---|---|---|---|---|---|---|
| *SOX17* | Forward | GGACCGCACGGAATTTGAAC | 60.11 | 55.00 | 21.00 | 127 | 109 |
| | Reverse | GGACACCACCGAGGAAATGG | 60.68 | 60.00 | 21.00 | | |
| *CDH5* | Forward | CCATTGAGACAAACCCCGCC | 61.60 | 60.00 | 20.00 | 138 | 84 |
| | Reverse | CGGGAGGGCTCATGTATCGG | 62.16 | 65.00 | 20.00 | | |
| *RUNX1* | Forward | GGTCGAAGTGGAAGAGGGAAAA | 60.22 | 50.00 | 22.00 | 61 | 117 |
| | Reverse | GCGGTGGGTTTGTGAAGACA | 61.10 | 55.00 | 20.00 | | |
| *GATA2* | Forward | CCTGTTGTGCAAATTGTCAGACG | 61.09 | 47.83 | 23.00 | 121 | 147 |
| | Reverse | GTCAGTGGCCTGTTAACATTGTG | 60.30 | 47.83 | 23.00 | | |
| *HOXA9* | Forward | ACAATGCTGAGAATGAGAGCGG | 61.00 | 50.00 | 23.00 | 65 | ND |
| | Reverse | GTTGGCTGCTGGGTTATTGGG | 61.84 | 57.14 | 21.00 | | |
| *SPINK2* | Forward | TCTGTGCATGAAAATCAGGGAA | 58.23 | 40.91 | 22.00 | 116 | 97 |
| | Reverse | TAGTCTGCCAGTGAAGGTGG | 59.02 | 55.00 | 20.00 | | |
| *RSPO3* | Forward | GAGTGTGTCAGTATTGTGCACTG | 59.81 | 47.83 | 23.00 | 106 | 126 |
| | Reverse | CCCGTGTTTCAGTCCCTCTTT | 60.20 | 52.38 | 21.00 | | |
| *ACTB* | Forward | GCCAACCGCGAGAAGATGAC | 61.70 | 60.00 | 20.00 | 87 | 135 |
| | Reverse | CAGGGATAGCACAGCCTGGA | 61.34 | 60.00 | 20.00 | | |
| *GAPDH* | Forward | GAGTGTGTCAGTATTGTGCACTG | 59.81 | 47.83 | 23.00 | 106 | 109 |
| | Reverse | CCCGTGTTTCAGTCCCTCTTT | 60.20 | 52.38 | 21.00 | | |

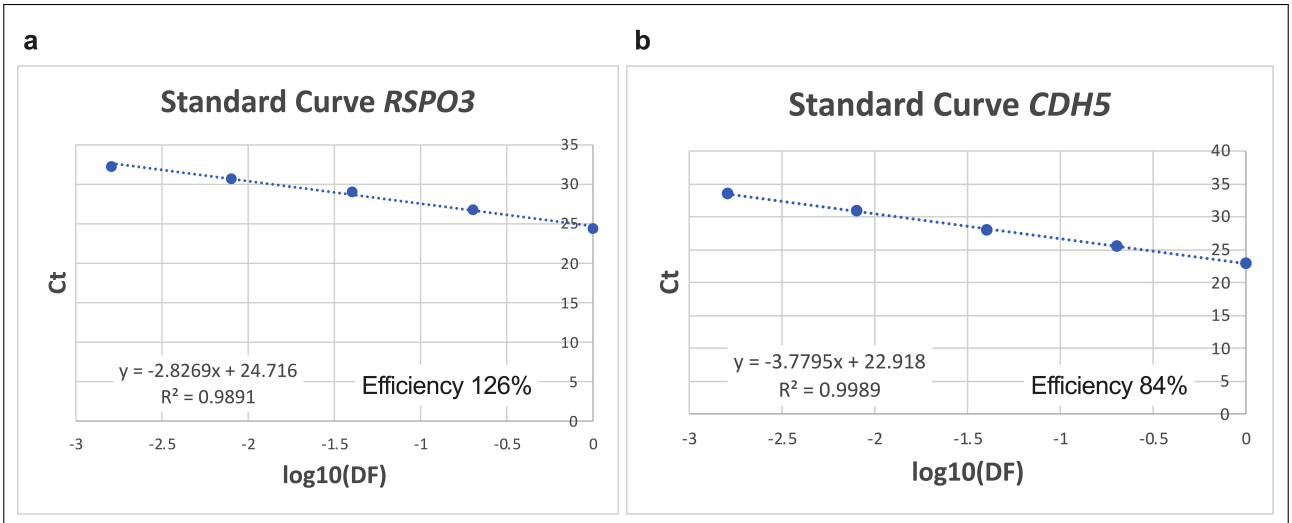


**Figure S1: Standard Curves from *RSPO3* and *CDH5* Optimization**

**a, b,** Standard curves for *RSPO3* and *CDH5*. The calculated amplification efficiencies are indicated in each plot.

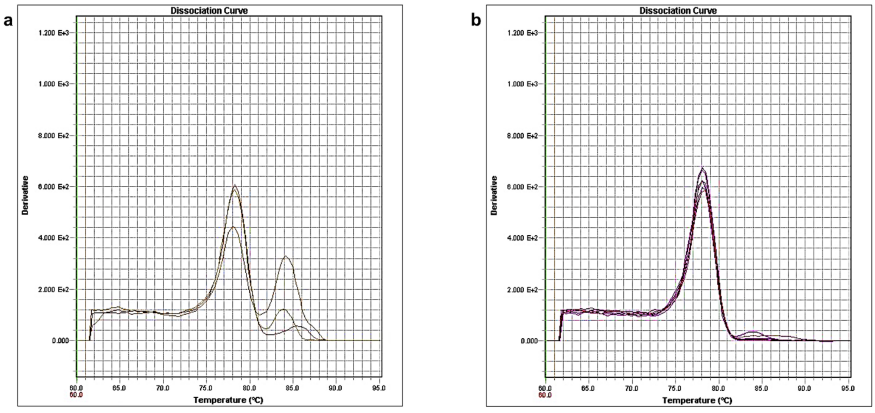


**Figure S2: Representative Dissociation Curves for Unspecific and Specific Amplification**

**a, b,** Plots showing the dissociation curve for *RUNX1*. **a,** The curves displaying two peaks indicate unspecific amplification. **b,** The uniform dissociation curves with the singular peak indicate specific amplification.

**Table A.3:** Table displaying antibodies used to stain samples for flow cytometry, antibody clones, company where the antibodies were acquired from, and final dilution factor

| Antibody Conjugate | Clone | Company | Dilution Factor |
|---|---|---|---|
| CD144 (VE-Cad) PerCP/Cy5.5 | 55-7H1 | BD Pharmingen | 1:50 |
| CD73 PE | AD2 | BD Pharmingen | 1:100 |
| CD43 APC/Cy7 | 1G10 | BD Biosciences | 1:50 |
| CD45 FITC | HI30 | eBioscience (ThermoFisher) | 1:50 |
| DAPI Solution | - | BD Biosciences | 1:100 |

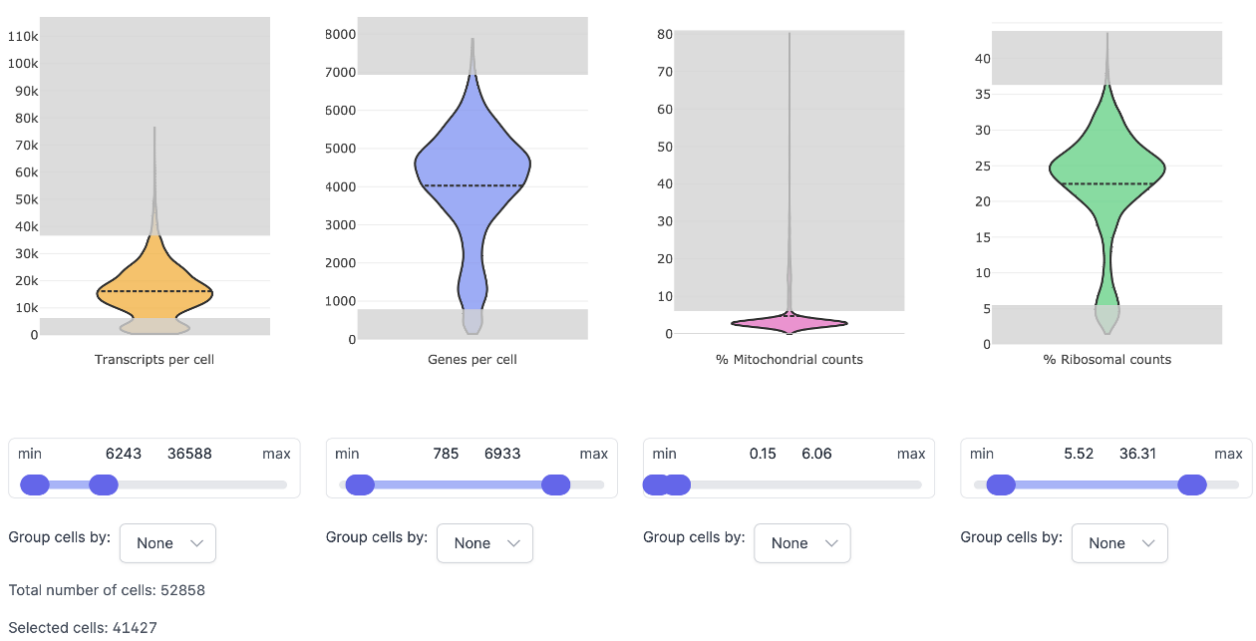


**Figure S3: Exclusion of Low Quality Cells**

Figure display the interface of the ScarfWeb online tool for excluding low quality cells. The violin plots illustrates the distributions of transcripts per cell, genes per cell, mitochondrial counts (%), and ribosomal counts (%), respectively. Low-quality cells are excluded based on thresholds set for each parameter, as indicated by the sliders. After filtering, 41,427 cells were selected for further analysis.

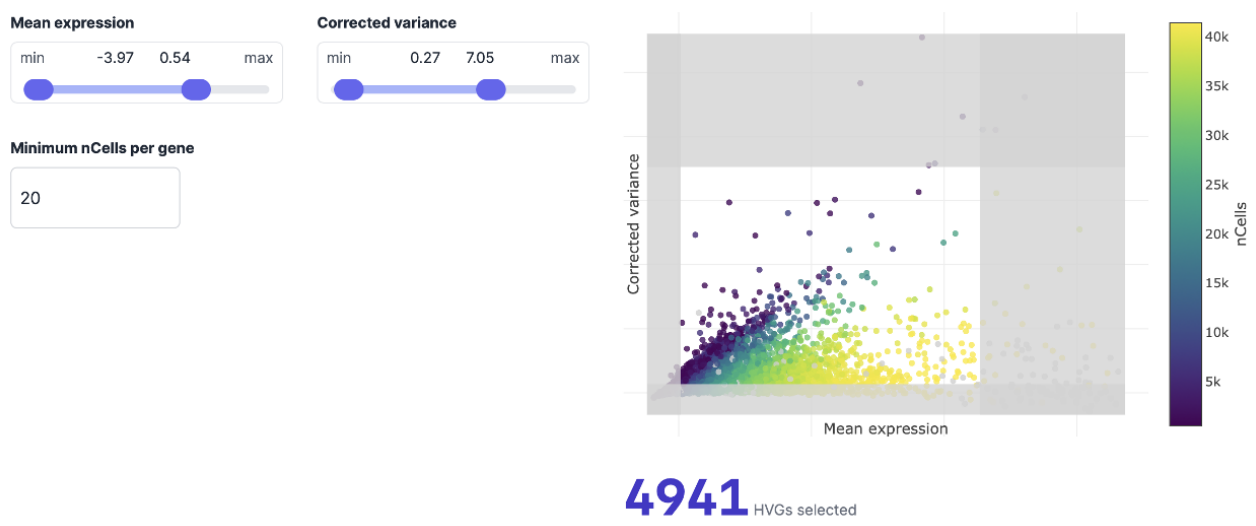


**Figure S4: Selection of Highly Variable Genes for UMAP Analysis**

Figure displays the interface of the ScarfWeb online tool for the selection of highly variable genes. The plot displays the relationship between mean expression and corrected variance of genes in a single-cell dataset, colored by the number of cells expressing each gene (nCells). Genes are selected as highly variable genes based on thresholds set for mean expression and variance, avoiding extremes to ensure robust data representation.

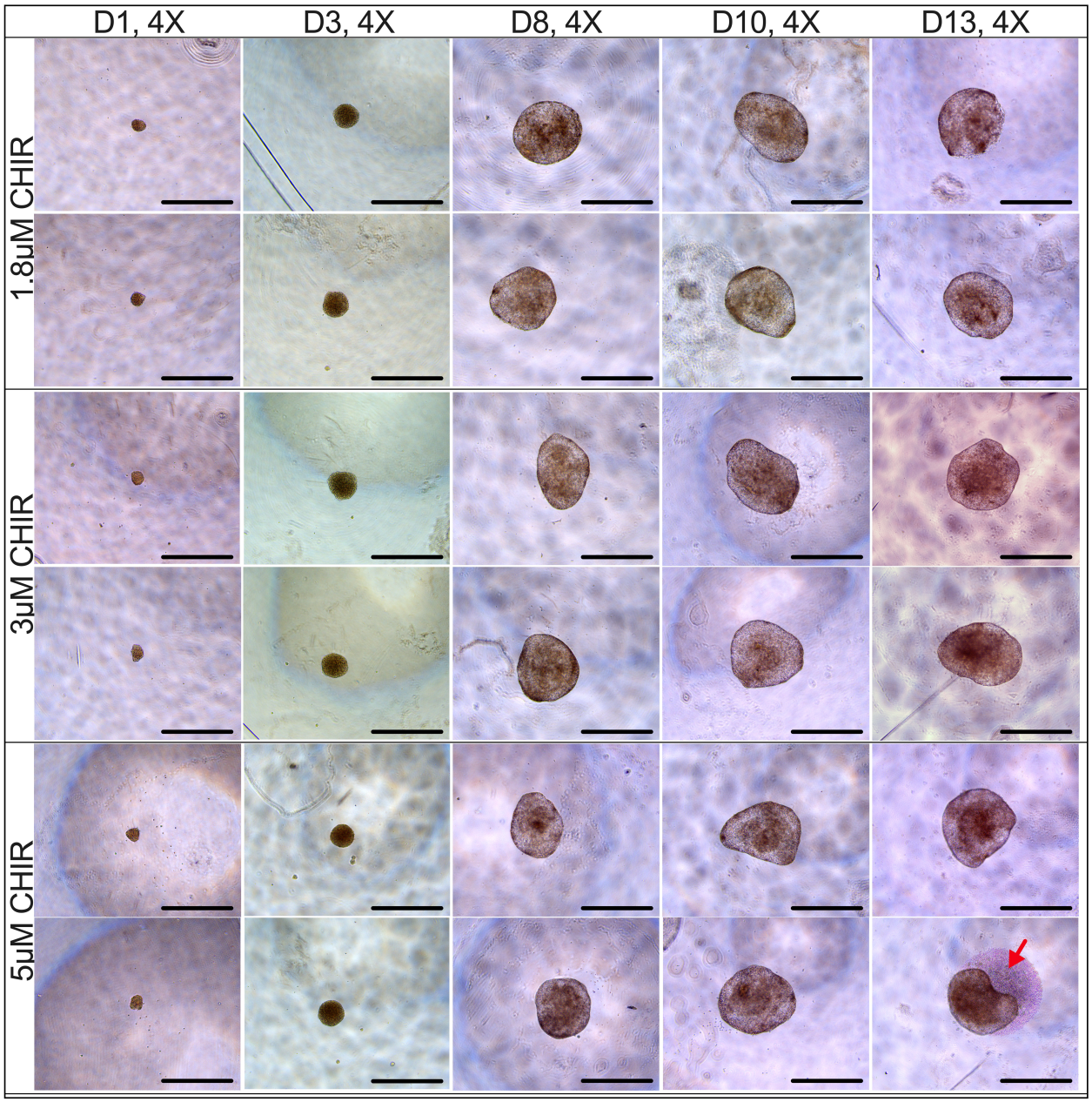


**Figure S5: Microscopy Images of Generated Embryoid Bodies for Each Condition**

Microscopy images at 4X magnification of developing hiPSC-derived EBs at Days 1, 3, 8, 10, and 13 post-EB formation (notated as D1, D3, D8, D10, and D13). Dispersed cells indicated by the red arrow. Note: Scale bars represent 1 mm.

**Table A.4:** Table displaying for each sample: concentration of CHIR used, volume RLT added when harvesting cells, total volume of the RNA sample after RLT addition, RNA quantification results (concentration and absorbances), kit used to synthesize cDNA, dilution factor applied to synthesized cDNA sample, and estimated cDNA concentration

| Experiment Timepoint.BioReplica | Condition | Volume RLT Added [µL] | Volume Total RNA Sample [µL] | RNA Concentration [ng/µL] | A(260/280) | A(260/230) | cDNA Synthesis Kit | Dilution Factor | Estimated cDNA Concentration [ng/µL] |
|---|---|---|---|---|---|---|---|---|---|
| EHB1 D0 | BPT | 350 | 400 | 143.70 | 1.99 | 1.95 | QuantiTect | 1:20 | 4.45 |
| EHB2 D0 | BPT | 350 | 400 | 234.00 | 1.99 | 2.07 | QuantiTect | 1:20 | 7.24 |
| EHB1 D0 | PT | 350 | 400 | 404.30 | 2.03 | 1.98 | QuantiTect | 1:20 | 12.51 |
| EHB2 D0 | PT | 350 | 400 | 127.40 | 1.97 | 1.54 | QuantiTect | 1:20 | 3.94 |
| EHB1 D3.1 | 1.8 µM CHIR | 75 | 200 | 120.40 | 1.98 | 1.38 | SuperScript III | 1:5 | 9.61 |
| EHB1 D3.2 | 1.8 µM CHIR | 75 | 230 | 70.40 | 2.00 | 1.82 | SuperScript III | 1:20 | 1.41 |
| EHB1 D3.1 | 3 µM CHIR | 75 | 130 | 212.10 | 1.99 | 1.91 | SuperScript III | 1:20 | 4.23 |
| EHB1 D3.2 | 3 µM CHIR | 75 | 200 | 20.40 | 1.90 | 1.31 | SuperScript III | 1:20 | 0.41 |
| EHB1 D3.1 | 5 µM CHIR | 75 | 220 | 47.40 | 1.78 | 1.33 | SuperScript III | 1:5 | 3.79 |
| EHB1 D3.2 | 5 µM CHIR | 75 | 230 | 140.80 | 1.99 | 1.82 | SuperScript III | 1:20 | 2.81 |
| EHB2 D3.1 | 1.8 µM CHIR | 75 | 180 | 220.70 | 2.05 | 1.56 | QuantiTect | 1:20 | 4.56 |
| EHB2 D3.2 | 1.8 µM CHIR | 75 | 180 | 177.50 | 2.06 | 0.74 | QuantiTect | 1:20 | 3.67 |
| EHB2 D3.1 | 3 µM CHIR | 75 | 180 | 196.50 | 2.08 | 0.67 | QuantiTect | 1:20 | 4.06 |
| EHB2 D3.2 | 3 µM CHIR | 75 | 180 | 220.50 | 2.06 | 1.68 | QuantiTect | 1:20 | 4.55 |
| EHB2 D3.1 | 5 µM CHIR | 75 | 180 | 152.10 | 2.00 | 2.05 | QuantiTect | 1:20 | 3.43 |
| EHB2 D3.2 | 5 µM CHIR | 75 | 180 | 159.60 | 2.02 | 2.12 | QuantiTect | 1:20 | 3.60 |
| EHB1 D8.1 | 1.8 µM CHIR | 75 | 120 | 120.50 | 1.99 | 1.98 | QuantiTect | 1:20 | 3.13 |
| EHB1 D8.2 | 1.8 µM CHIR | 75 | 120 | 65.70 | 1.96 | 1.10 | QuantiTect | 1:20 | 2.03 |
| EHB1 D8.1 | 3 µM CHIR | 75 | 110 | 39.50 | 2.03 | 0.96 | SuperScript III | 1:20 | 0.79 |
| EHB1 D8.2 | 3 µM CHIR | 75 | 110 | 53.80 | 2.06 | 1.45 | SuperScript III | 1:20 | 1.07 |
| EHB1 D8.1 | 5 µM CHIR | 75 | 120 | 45.90 | 2.02 | 1.09 | QuantiTect | 1:20 | 1.42 |
| EHB1 D8.2 | 5 µM CHIR | 75 | 120 | 45.70 | 2.08 | 0.09 | QuantiTect | 1:20 | 1.41 |
| EHB2 D8.1 | 1.8 µM CHIR | 75 | 105 | 143.90 | 2.03 | 0.75 | QuantiTect | 1:20 | 3.35 |
| EHB2 D8.2 | 1.8 µM CHIR | 75 | 105 | 100.10 | 1.97 | 2.05 | QuantiTect | 1:20 | 2.78 |
| EHB2 D8.1 | 3 µM CHIR | 75 | 105 | 119.10 | 1.98 | 1.14 | QuantiTect | 1:20 | 3.31 |

Continued on next page

Table A.4 – Continued from previous page

| Experiment Timepoint.BioReplica | Condition | Volume RLT Added [µL] | Volume Total RNA Sample [µL] | RNA Concentration [ng/µL] | A(260/280) | A(260/230) | cDNA Synthesis Kit | Dilution Factor | Estimated cDNA Concentration [ng/µL] |
|---|---|---|---|---|---|---|---|---|---|
| EHB2 D8.2 | 3 µM CHIR | 75 | 105 | 148.20 | 2.00 | 1.59 | QuantiTect | 1:20 | 3.45 |
| EHB2 D8.1 | 5 µM CHIR | 75 | 105 | 85.50 | 2.07 | 0.14 | QuantiTect | 1:20 | 2.65 |
| EHB2 D8.2 | 5 µM CHIR | 75 | 105 | 80.70 | 1.94 | 0.55 | QuantiTect | 1:20 | 2.50 |
| EHB1 D13.1 | 1.8 µM CHIR | 350 | 400 | 22.70 | 1.85 | 1.42 | QuantiTect | 1:5 | 1.60 |
| EHB1 D13.2 | 1.8 µM CHIR | 350 | 400 | 14.20 | 2.08 | 0.55 | QuantiTect | 1:5 | 1.00 |
| EHB1 D13.1 | 3 µM CHIR | 350 | 400 | 31.90 | 1.97 | 1.35 | QuantiTect | 1:20 | 0.56 |
| EHB1 D13.2 | 3 µM CHIR | 350 | 400 | 41.40 | 1.95 | 1.83 | QuantiTect | 1:20 | 0.73 |
| EHB1 D13.1 | 5 µM CHIR | 350 | 400 | 40.40 | 1.93 | 1.18 | QuantiTect | 1:20 | 0.71 |
| EHB1 D13.2 | 5 µM CHIR | 350 | 400 | 15.70 | 1.83 | 1.10 | QuantiTect | 1:5 | 1.10 |
| EHB2 D13.1 | 1.8 µM CHIR | 350 | 400 | 30.10 | 1.89 | 1.35 | QuantiTect | 1:20 | 0.53 |
| EHB2 D13.2 | 1.8 µM CHIR | 350 | 400 | 36.50 | 1.96 | 1.78 | QuantiTect | 1:20 | 0.64 |
| EHB2 D13.1 | 3 µM CHIR | 350 | 400 | 74.20 | 2.00 | 1.91 | QuantiTect | 1:20 | 1.30 |
| EHB2 D13.2 | 3 µM CHIR | 350 | 400 | 76.10 | 1.98 | 1.82 | QuantiTect | 1:20 | 1.34 |
| EHB2 D13.1 | 5 µM CHIR | 350 | 400 | 206.30 | 2.06 | 2.04 | QuantiTect | 1:20 | 1.74 |
| EHB2 D13.2 | 5 µM CHIR | 350 | 400 | 114.90 | 2.03 | 1.73 | QuantiTect | 1:20 | 1.66 |

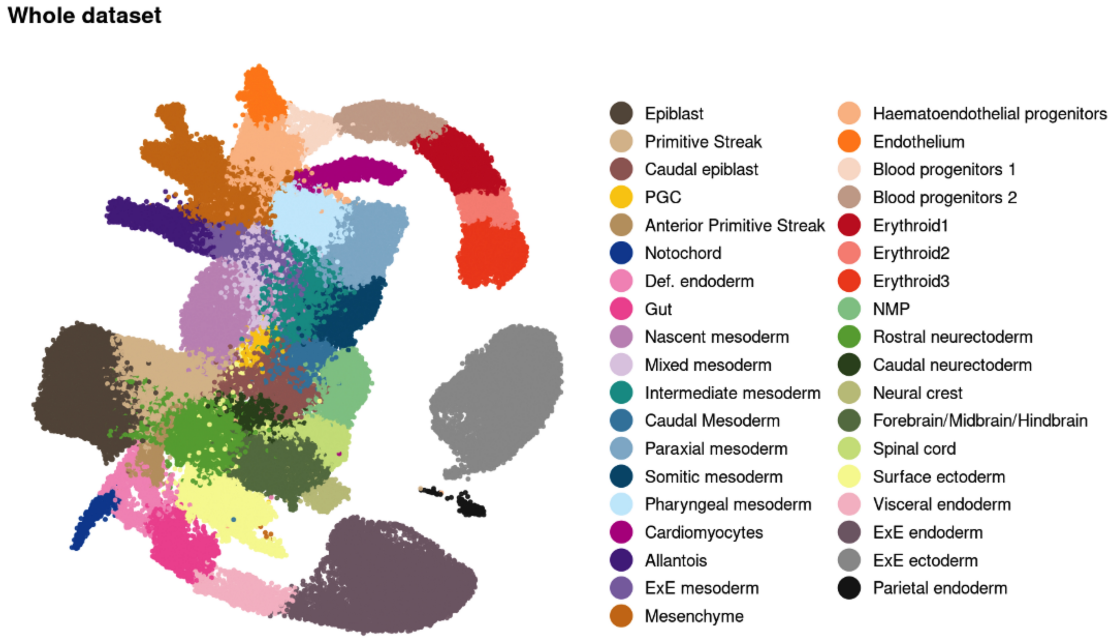


**Figure S6: Complete UMAP of Gastrulation and Early Organogenesis**

UMAP plot of gastrulation and early organogenesis [58] mapping 116,312 cells. Cells are colored by their annotated cell type. ExE: Extra-embryonic, NMP: Neuromesodermal progenitors, PGC: Primordial germ cells, Def.: Definitive.

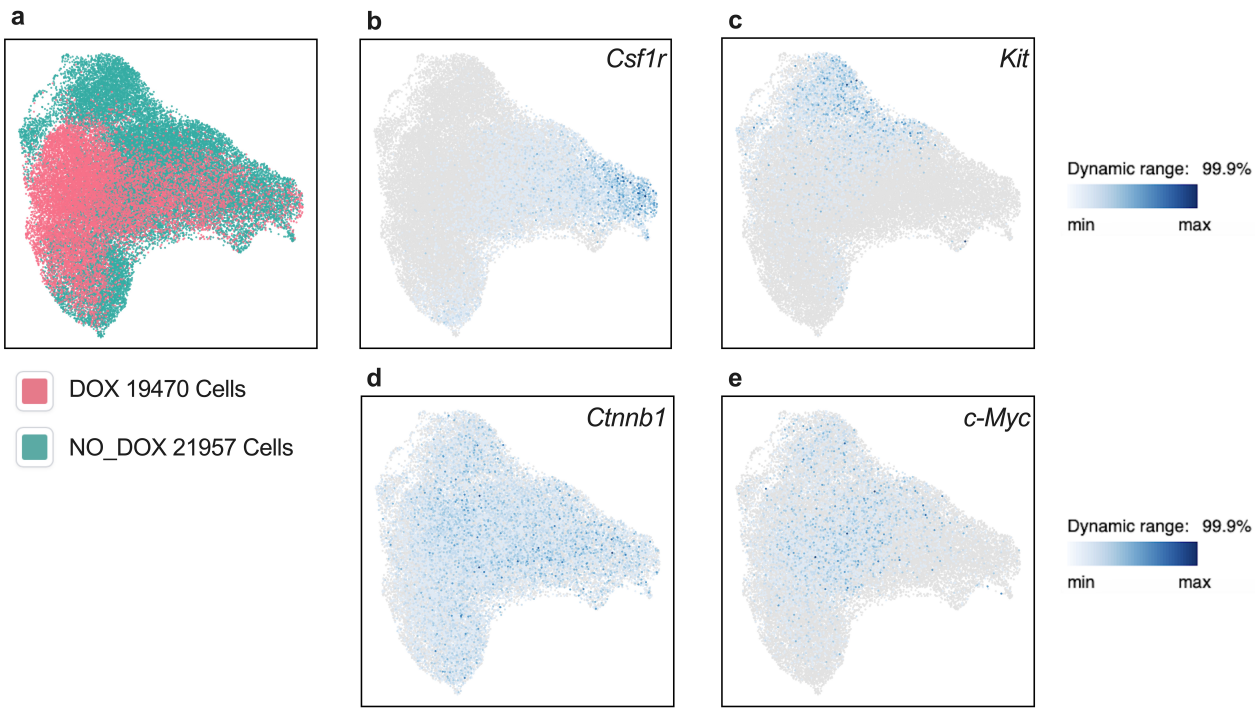


**Figure S7: Various UMAP Plots of the AML Dataset**

**a** UMAP plot of leukemia dataset mapping a total of 41,427 cells. Cells are clustered by conditions, DOX or NO DOX. **b-e** UMAPs visualizing the analysed leukemia dataset, with cells colored by the expression of *Csf1r*, *Kit*, *Ctnnb1*, and *c-Myc*, respectively. *Csf1r* is a marker for mature hematopoietic cells toward the myeloid lineage, while *Kit* is related to immature cells such as HSCs. *Ctnnb1* and *c-Myc* are markers for the Wnt signaling pathway. Color intensity indicates expression levels, with darker shades of blue representing higher expression and grey color representing absent expression.

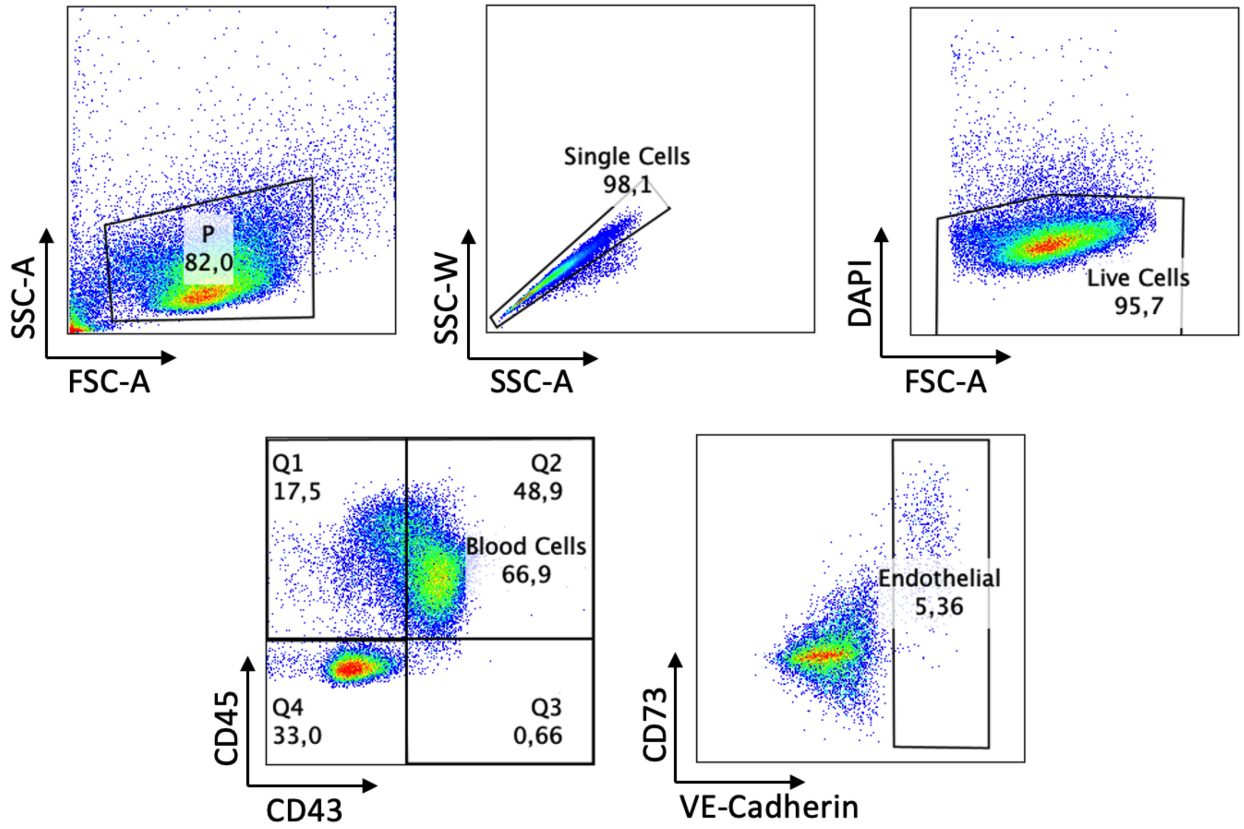


**Figure S8: Representative Flow Cytometry Gating Strategy**

Flow cytometry gating strategy using the 3 µM condition at D13 (Exp EHB2) as reference. Scattering settings and markers used as indicated by the plot axes. The first gate includes single cells. Within this gate the live cells are gated for. Within the live cell population, the blood cells are marked by CD45 and CD43. Within the CD45- CD43- quadrant the endothelial cell population is marked by VE-Cadherin. Percentages represent frequencies of cell populations within parent popuation.

**DEPARTMENT OF MATHEMATICAL SCIENCES**
**CHALMERS UNIVERSITY OF TECHNOLOGY**
Gothenburg, Sweden
www.chalmers.se


CHALMERS
UNIVERSITY OF TECHNOLOGY