

## REVIEW ARTICLE OPEN



# Exploring hematopoiesis in zebrafish using forward genetic screening

Hyemin Song<sup>1,2</sup>, Unbeom Shin<sup>3</sup>, Uijeong Nam<sup>4</sup> and Yoonsung Lee<sup>5</sup>✉

© The Author(s) 2023

Zebrafish have emerged as a powerful animal model for investigating the genetic basis of hematopoiesis. Owing to its close genetic and developmental similarities to humans, combined with its rapid reproduction and extensive genomic resources, zebrafish have become a versatile and efficient platform for genetic studies. In particular, the forward genetic screening approach has enabled the unbiased identification of novel genes and pathways related to blood development, from hematopoietic stem cell formation to terminal differentiation. Recent advances in mutant gene mapping have further expanded the scope of forward genetic screening, facilitating the identification of previously unknown genes and pathways relevant to hematopoiesis. In this review, we provide an overview of the zebrafish forward screening approach for hematopoietic gene discovery and highlight the key genes and pathways identified using this method. This review emphasizes the importance of zebrafish as a model system for understanding the genetic basis of hematopoiesis and its associated disorders.

*Experimental & Molecular Medicine* (2024) 56:51–58; <https://doi.org/10.1038/s12276-023-01138-2>

## INTRODUCTION

The zebrafish (*Danio rerio*) has been used as a model organism for studying hematopoiesis since the 1960s because of its experimental advantages, including high fecundity, optical transparency of embryos, and diverse gene manipulation techniques<sup>1</sup>. In addition, zebrafish possess evolutionarily conserved hematopoietic genes and regulatory networks, with a comprehensive characterization of all mature blood lineages and intermediates emphasizing its significance in this field<sup>2</sup>. Utilizing diverse genetic approaches in zebrafish, multiple pathways and molecular mechanisms that underlie hematopoiesis have been defined<sup>3</sup>.

Various gene editing approaches have been employed to elucidate the genes associated with hematopoietic processes using zebrafish. The knockdown approach, commonly performed by morpholinos, provides a convenient mechanism to suppress gene expression, allowing rapid insights into gene function<sup>4</sup>. However, this approach has limitations due to concerns about possible off-target effects and inconsistent knockdown efficiencies<sup>5</sup>.

Reverse genetic screening techniques, exemplified by transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9, enable the precise targeting and modification of specific genes, serving as powerful tools for validation<sup>6,7</sup>. Notably, the CRISPR system, utilizing guide RNA (gRNA) and the Cas9 enzyme, facilitates efficient gene targeting in terms of cost and time<sup>8</sup>. Over the past decade, the zebrafish model has harnessed the CRISPR system not only to produce knockout (KO) animals but also to execute precise mutagenesis through techniques such as base

editing and prime editing<sup>9–11</sup>. While challenges such as off-target effects persist, continuous advancements are being made in these gene-editing techniques.

In addition to reverse genetic methodologies, forward genetic screening is another hallmark of zebrafish research. This unbiased approach provides a platform for identifying novel genes critical for hematopoiesis without predetermined notions. Large-scale forward genetic screening allows the identification of unsuspected players in hematopoiesis, offering insights into the intricate regulatory networks governing this process<sup>12,13</sup>.

To perform a general forward genetic screening in zebrafish, random point mutations are initially induced throughout the genome of adult male zebrafish using a chemical mutagen<sup>14,15</sup>. Inbreeding of heterozygous progeny generated from mutagenized zebrafish produces homozygous mutations in the F3 generation. Through various phenotypic characterizations of hematopoietic events, numerous mutants with hematopoietic defects have been obtained<sup>12,16</sup>. Genetic linkage mapping, along with known polymorphic repeat markers, is critical for identifying the candidate genes responsible for the phenotypes of interest. However, the identification of these causative mutations has been severely limited owing to insufficient genetic markers and laborious work demands<sup>14,17</sup>.

Targeting Induced Local Lesions in Genomes (TILLING) is a hybrid technique that combines elements from both forward and reverse genetic screening. This method involves examining genomic DNA samples from ENU-mutagenized zebrafish to detect mutations in a selected gene. Using sperm from the respective animal, the mutant line is re-established for phenotype

<sup>1</sup>Department of Biomedical Sciences, UC San Diego School of Medicine, La Jolla, CA 92093, USA. <sup>2</sup>Salk Institute for Biological Studies, La Jolla, CA 92037, USA. <sup>3</sup>School of Life Sciences, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea. <sup>4</sup>Department of Biomedical Science and Technology, Kyung Hee University, Seoul 05278, Republic of Korea. <sup>5</sup>Clinical Research Institute, Kyung Hee University Hospital at Gangdong, School of Medicine, Kyung Hee University, Seoul 05278, Republic of Korea. ✉email: ylee3699@khu.ac.kr

Received: 14 June 2023 Revised: 11 October 2023 Accepted: 13 October 2023

Published online: 4 January 2024

characterization<sup>18</sup>. TILLING employs the advantages of forward genetics while also facilitating targeted analysis of specific candidate genes, resulting in the identification of over 150 loss-of-function mutations<sup>19</sup>. Although TILLING provides powerful genetic capabilities, it can be resource intensive, time-consuming, and relatively less effective at inducing complete KOs.

Recently, the advent of next-generation sequencing (NGS) has introduced new genomic and transcriptomic approaches that can sequence nearly the entire zebrafish genome of interest<sup>20,21</sup>. Using a bioinformatics pipeline, thousands of markers across almost any genome of interest can be used to identify causative mutations found by forward screening in a fast, simple, and reliable manner.

In this review, we discuss the application of forward genetic screening in zebrafish to identify significant genes and pathways related to hematopoiesis. We also summarize the key genes discovered through forward genetic approaches and technologies that have helped to dissect the molecular mechanisms of genes involved in hematopoiesis, especially in hematopoietic stem cell (HSC) development. This review provides a comprehensive overview of our current understanding of the molecular mechanisms underlying the development of the hematopoietic system, with a particular focus on zebrafish as a powerful genetic model organism.

## ZEBRAFISH HEMATOPOIESIS

The genetic and regulatory networks of hematopoiesis in zebrafish resemble those in humans. In zebrafish, similar to other vertebrates, hematopoiesis occurs in two distinct waves: primitive and definitive hematopoiesis<sup>22</sup>. Primitive hematopoiesis produces primitive erythroid and myeloid cells by differentiating hemangioblasts, which give rise to both blood cells and vessels<sup>3,23</sup>. It mainly occurs in the anterior and posterior lateral mesoderm, which later becomes the intermediate cell mass where HSCs are congregated (Fig. 1a)<sup>22,24,25</sup>. The generated primitive erythroid and myeloid cells play a crucial role in the transfer of oxygen to tissues and immunity of early embryos and start to circulate throughout the embryos at approximately 24 h post-fertilization (hpf)<sup>26,27</sup>.

Definitive HSCs, which are pluripotent stem cells that give rise to all blood cell lineages, begin to emerge from the ventral wall of the dorsal aorta (DA), a region known as the aorta-gonad-mesonephros, beginning approximately 30 hpf (Fig. 1b)<sup>28–31</sup>. Subsequently, at 2 days post-fertilization (dpf), these cells start migrating toward the caudal hematopoietic tissue (CHT), where they mature and continuously circulate before finally settling in the kidney (Fig. 1b)<sup>32,33</sup>. These definitive HSCs reside in the kidney marrow, which is analogous to the mammalian bone marrow, and produce all the blood lineages necessary for normal blood function from embryonic development to adulthood (Fig. 1c)<sup>22,30,34</sup>. Although many genes and pathways responsible for HSC formation and differentiation have been identified, much remains to be understood. To study the precise steps of HSC specification, differentiation, and maintenance, researchers have utilized zebrafish mutant models generated through various mutagenesis techniques.

## FORWARD GENETIC SCREENING FOR ZEBRAFISH HEMATOPOIESIS

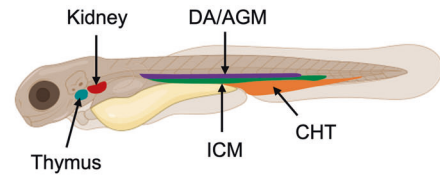
### N-ethyl-N-nitrosourea (ENU) random mutagenesis

Zebrafish have become a valuable model for large-scale genetic screening because of their ease of genetic manipulation. They are particularly suitable for the unbiased forward screening of specific phenotypes, which is almost exclusively practical among vertebrate models<sup>35</sup>. In forward genetics, random mutations are induced in an organism, and the resulting phenotypes are analyzed to identify mutations that affect the process of interest. Random mutagenesis can be achieved through chemical mutagens such as ENU, physical mutagens such as ionizing radiation, or transposon mutagenesis, all of which randomly induce DNA damage. Among these methods,

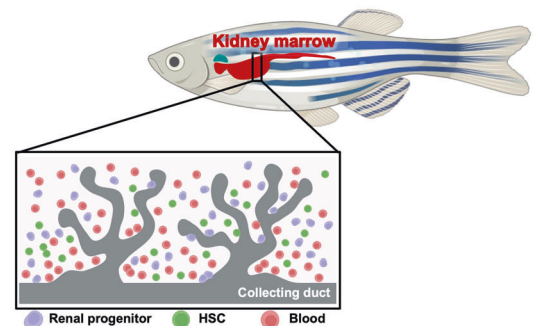
### a Primitive hematopoiesis



### b Definitive hematopoiesis



### c Adult hematopoiesis

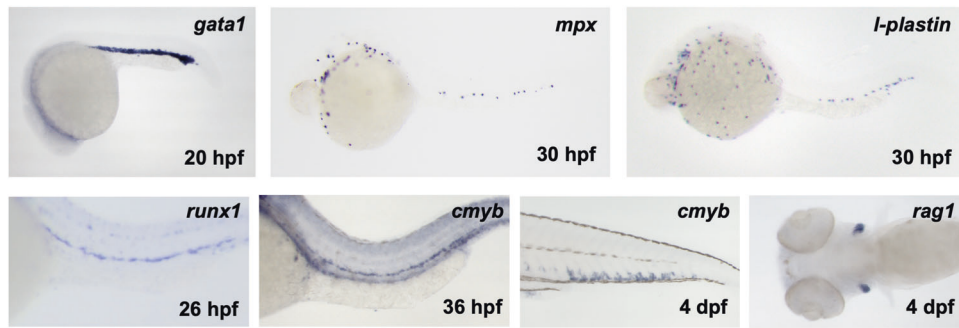


**Fig. 1 Schematic representation of zebrafish hematopoiesis.** **a** Primitive hematopoiesis takes place in two locations, the anterior lateral mesoderm (green) and posterior lateral mesoderm (red), both of which are critical for blood and endothelial formation during the somitogenesis stages. **b** Definitive hematopoietic stem cells (HSCs) arise from the ventral wall of the dorsal aorta (DA) and migrate to the caudal hematopoietic tissue (CHT). HSCs then migrate from the CHT to the kidney, where they undergo lifelong hematopoiesis. HSCs then seed the thymus, where they mature and differentiate into T lymphocytes. **c** Definitive HSCs and other blood cells are distributed within the kidney marrow of adult zebrafish. AGM, aorta gonad mesonephros; ICM, intermediate cellular mass. Figure generated with [BioRender.com](https://www.biorender.com).

ENU mutagenesis is a common and efficient chemical mutagenesis method in zebrafish that can induce point mutations randomly in germ cells<sup>36</sup>. In ENU mutagenesis, adult male fish are exposed to water containing ENU for approximately one hour per week over the course of three weeks. The use of anesthetics can increase the survival rate of male fish during this process. ENU mutagenesis induces random single base pair mutations with a high success rate for mutagenic loads in zebrafish premeiotic germ cells<sup>17,37</sup>. Mutagenized zebrafish can exhibit phenotypes that are discretely linked to mutated lesions in a single gene.

### Phenotype screening for zebrafish hematopoiesis

Early zebrafish forward screens primarily relied on visually observing live embryos to detect specific morphological defects<sup>38</sup>. Currently, to identify mutants with hematopoietic deficiencies, physically collected mutants displaying distinct morphological defects can be screened for blood development using molecular markers for a given cell or tissue type. Whole-mount in situ hybridization (WISH) is a simple and commonly used approach for detecting the expression site of genes of interest using digoxigenin-labeled antisense RNA probes<sup>39</sup>. Using various hematopoietic markers, the emergence and



**Fig. 2 Whole-mount in situ hybridization (WISH) readouts labeling primitive and definitive hematopoiesis.** Representative images of WISH using probes labeling primitive erythroid (*gata1*), primitive myeloid (*mpx*, *l-plastin*), definitive HSCs (*runx1*, *cmyb*), and thymocytes (*rag1*) at distinct developmental stages.

differentiation of HSCs and other leukocytes can be easily examined in mutant embryos, which can elucidate whether the mutants have hematopoietic developmental defects<sup>40,41</sup>. Moreover, WISH is a useful tool for exploring various tissues and cells important for hematopoiesis, such as primitive erythroid and myeloid cells, as well as definitive HSCs and lymphoid cells (Fig. 2)<sup>42</sup>. Additionally, immunostaining using specific antibodies to observe mutant phenotypes can visualize the development of tissues essential for hematopoiesis processes<sup>41</sup>.

Although WISH and immunostaining are beneficial for increasing the range of phenotypes that can be identified in mutants, these methods require sample fixation, which limits subsequent observations. Tissue-specific transgenic lines have been developed in zebrafish to mark specific cell lineages with fluorescent proteins, enabling the visualization of circulating blood cells in live animals<sup>43,44</sup>. Real-time monitoring of leukocytes facilitates the identification of hematopoietic defects in mutant zebrafish. Large-scale screening using transgenic lines is a feasible approach to assess the development of distinct blood cells and other tissues essential for hematopoiesis. By marking multiple cells and tissues with different fluorescent proteins, transgenic lines facilitate the simultaneous monitoring of the development of blood cells and other hematopoietic tissues<sup>45,46</sup>. Additionally, imaging with transgenic lines yields higher-resolution images compared to those obtained through immunostaining and WISH, leading to more reliable analyses.

## LINKAGE GENETIC MAPPING THROUGH NGS

### Traditional linkage analysis

After generating zebrafish mutants and determining their phenotypes through forward screening, the causative genes responsible for the observed phenotype should be identified to further understand the underlying genetic mechanisms. Linkage mapping is a genetic approach used to identify the genomic region responsible for a specific phenotype of interest in mutants<sup>41</sup>. This method involves using known genetic markers on zebrafish chromosomes to map the phenotypes to specific regions of the genome. Genetic markers include a variety of forms, including restriction fragment length polymorphisms, simple sequence length polymorphisms (SSLPs), insertion–deletion polymorphisms and single-nucleotide polymorphisms (SNPs). These markers are used to track inheritance patterns and map the locations of genes of interest on the chromosome. Once the phenotype of the mutant is linked to a specific region of the genome, the genes responsible for the observed phenotypes can be identified by sequencing the genes in the identified genomic region of the mutants and recovering the altered phenotypes of the mutants by over-expressing the normal gene products of the mutated genes. Although traditional linkage mapping is a powerful tool for identifying novel genes associated with mutants generated via

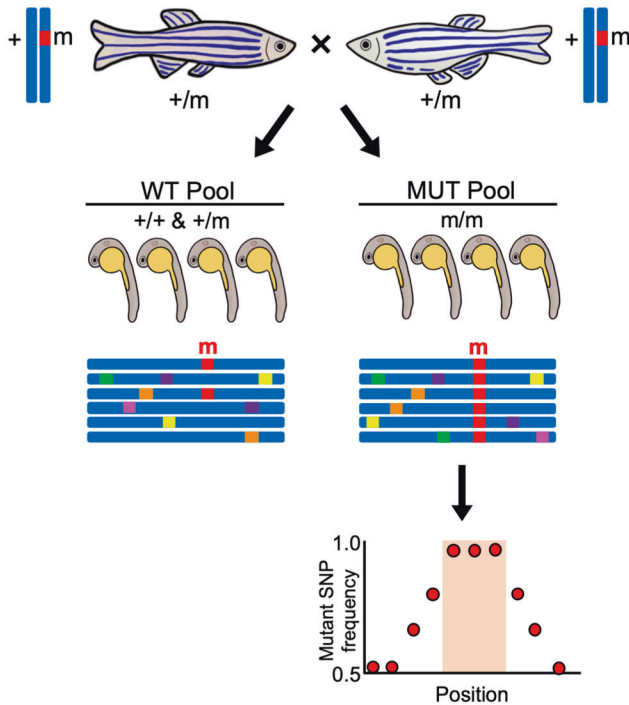
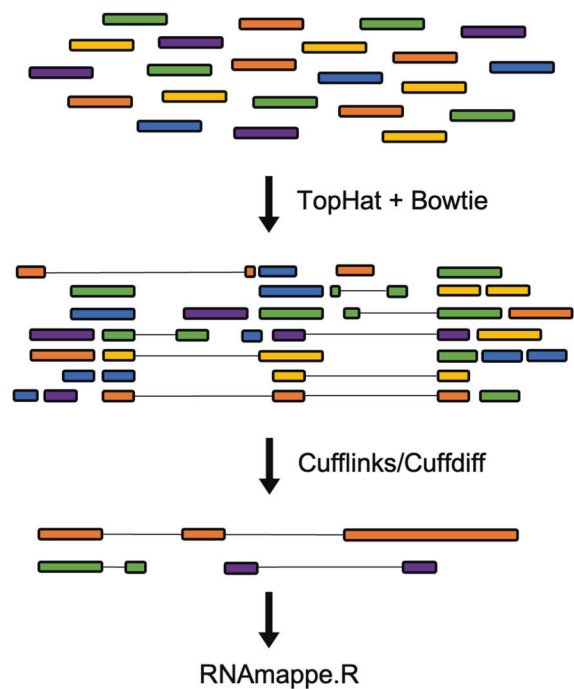
forward genetics, it has several limitations. This method is time-consuming and requires significant resources and expertise. Due to these difficulties, in the first large-scale zebrafish forward genetics study, only 30% of the genes in the mutants could be identified using traditional linkage mapping<sup>14</sup>.

### Whole-genome sequencing (WGS)-oriented linkage mapping

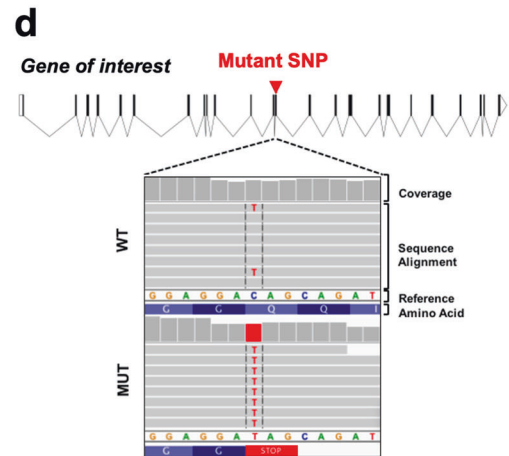
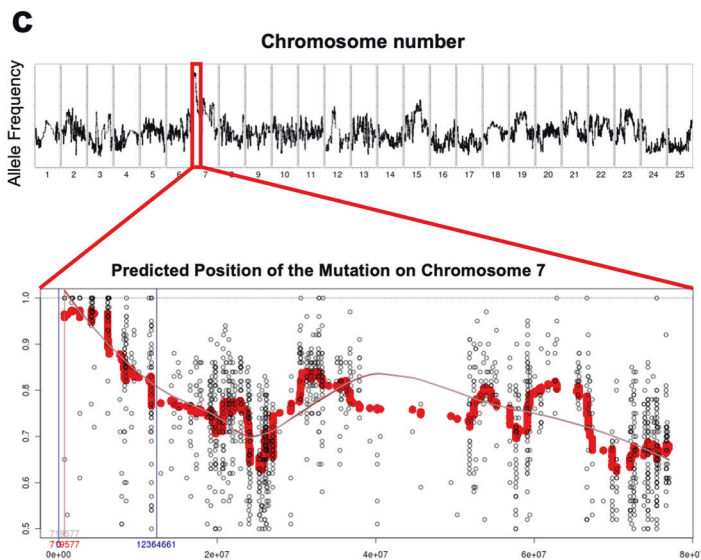
Advances in genetics have led to the development of molecular tools that enable genome-wide analyses to identify the genomic region linked to a mutation. More recently, NGS has become an innovative method for identifying putative genes that may cause the phenotype of interest in zebrafish whole-genome studies<sup>15,47,48</sup>. Although WGS has been widely used to identify candidate genes in *Caenorhabditis elegans* and *Drosophila*<sup>49,50</sup>, it is not a particularly efficient method for zebrafish because of their large and highly polymorphic genome<sup>51</sup>. Thus, sequencing of a single zebrafish genome using WGS is inadequate to accurately distinguish potential causative mutations from other polymorphisms. Instead, bioinformatics approaches have been applied to improve the reliability of sequencing and genetic mapping methods. However, it does not directly address the issue of cost-effectiveness frequently associated with WGS.

To enhance the process of identifying candidate genes from forward genetic screening, RNA-seq-based bulk segregant analysis has been introduced<sup>52</sup>. This approach is not only cost-effective but also capable of mapping and identifying both genomic regions linked to mutations of interest and mutant genes. A bioinformatics pipeline that employs RNA-seq-based mapping has been developed to identify numerous genes associated with forward genetics. After conducting forward genetic screening, pools of wild-type and mutant zebrafish are separated according to their phenotypes, and their total RNA is processed to prepare RNA-seq samples (Fig. 3a). Following RNA extraction, sequencing libraries are generated for each pool, resulting in multiple samples from several different mutants during sequencing. To analyze the differential expression, RNA sequencing reads from each wild-type or mutant sample are independently aligned and assembled using the TopHat suite<sup>53</sup>. Differential expression is determined, and differentially expressed genes are identified using Cufflinks and the DEGseq R package (Fig. 3b)<sup>54</sup>.

The RNAmapper pipeline is used to input the aligned RNA sequencing datasets and identify the genomic region linked to the mutant phenotype using SNP markers<sup>52</sup>. When forward genetics yields a recessive mutation, the region linked to the mutation exhibits homozygosity. Consequently, a significant peak, indicating a high frequency of the mutant allele at and near the mutation site close to 1, is detected in the linked position of specific chromosomes (Fig. 3c). Once a linkage region is identified, all SNPs within the region are determined from the RNA-seq data of the mutant for candidate SNP filtering. The Ensembl Variant Effect Predictor tool is used to verify whether these candidate SNPs

**a** Forward Genetic Screen Strategy**b** TopHat/Cufflinks Workflow

## Candidate Identification through Linkage Analysis



**Fig. 3** Next-generation sequencing (NGS)-based linkage mapping. **a** Diagram of the forward genetic screening approach. **b** Workflow for transcriptome analysis and identification of differentially expressed genes using TopHat suite and Cufflinks. **c** Alignment of RNA-seq reads to the reference genome using the RNA mapper pipeline, which allows visualization of the whole genome as well as the linked interval on a specific chromosome with high homozygous allele frequency. **d** Position and RNA-seq coverage of SNPs in the gene of interest, which resulted in the introduction of a premature stop codon.

result in nonsynonymous changes<sup>55</sup>. Using the Integrative Genomics Viewer, each SNP is visually observed to ensure that the wild-type and mutant SNPs are segregated appropriately (Fig. 3d)<sup>56</sup>. Ideally, the putative mutant SNP is expected to be present in 100% of the mutants and 33% of the sibling wild-type embryos because of the presence of heterozygous siblings in the

pool. Further confirmation is performed using Sanger sequencing to ensure that the candidate SNPs are present in all the mutants.

Overall, RNA-seq-based linkage mapping greatly enhances the identification of a selected number of high-priority candidate mutations associated with phenotypic traits within the linked region. The integration of computational methods has further accelerated



**Table 1.** Genes involved in hematopoiesis identified using forward genetic screening.

Hematopoietic process	Mutant name	Gene	Phenotype	Readout for screening	Biological character of the mutants	Refs.
Primitive erythropoiesis	<i>moonshine</i>	<i>tif1r</i>	Bloodless	<i>gata1</i> WISH	Apoptosis in erythroid progenitor cells	12,58
	<i>vlad tepes</i>	<i>gata1</i>	Bloodless	Blood cell count	Inability to bind to the promoter of its target gene	16,59
	<i>t21384</i>	<i>tal1</i>	Vascular defects	Endothelial alkaline phosphatase activity	Involved in the differentiation of red blood cells	61,62
Primitive myelopoiesis	<i>mummy</i>	<i>dhx8</i>	No blood circulation	<i>cbfb</i> and <i>l-plastin</i> WISH	Affects mRNA splicing and cell division during mitosis	64
	<i>moli</i>	<i>cebpa</i>	Decreased macrophage and neutrophils	Neutral red, Sudan black B staining	Cell cycle arrest of myeloid progenitors	57
Early endothelial and blood cells	<i>cloche</i>	<i>PAS-domain-containing bHLH transcription factor</i>	Blood reduction and vascular defects	GATA-2 expression	Defects in the development of the heart, vasculature and blood	35,67
	<i>hi2335</i>	<i>plcg1</i>	Bloodless	<i>cmyb</i> WISH	Lack of artery specification	68
	<i>hi3714</i>	<i>tbx16</i>	Bloodless	<i>cmyb</i> WISH	Lack of axial vascular organization	68
	<i>hi1618</i>	<i>hdac1</i>	Bloodless	<i>cmyb</i> WISH	Defect in vessel or artery formation	68
HSC specification	<i>sd45</i>	<i>supt16h</i>	Lack of HSCs	<i>runx1</i> WISH	Enlarged hindbrain and curved tail	69
HSC maintenance	<i>rumba</i>	C2H2 zinc finger protein	Lack of T lymphocytes	<i>rag1</i> WISH	Delayed cell cycle	70
	<i>samba</i>	<i>has3</i>	Lack of T lymphocytes	<i>rag1</i> WISH	Cell arrested in M phase	70
	<i>lack</i>	<i>sae1</i>	Lack of T lymphocytes	<i>rag1</i> WISH	Impaired SUMOylation	71
	<i>116</i>	<i>rpc9</i>	T-cell deficient	<i>rag1</i> WISH	p53-dependent apoptosis	72
	<i>ceylon</i>	<i>tbl3</i>	Lack of lymphoid cells	<i>rag1</i> WISH	Reduced population of T cells and red blood cells	40,73
	<i>grechetto</i>	<i>cpsf1</i>	Bloodless	<i>mpx</i> WISH	Abnormal polyadenylation of gene needed for HSC development	74
	<i>cas003</i>	<i>topbp1</i>	Decreased HSCs	<i>cmyb</i> WISH	p53-dependent apoptosis due to abnormal ATR/Chk1 activation	75
	<i>cas002</i>	<i>kri11</i>	Decreased HSCs	<i>cmyb</i> WISH	Autophagy due to accumulation of unfolded protein activation of PERK	76
	<i>cas008</i>	<i>gemin5</i>	Reduced <i>myb</i> expression	<i>cmyb</i> WISH	Reduced proliferation of HSCs	77
	<i>smu07</i>	<i>slc20a1b</i>	Bloodless	<i>βe1-globin</i> WISH	Cell arrested in G2/M phase	78,79
HSC niche	<i>cas005</i>	<i>itga4</i>	Decreased HSCs	<i>cmyb</i> WISH	Defective interaction between HSC and macrophages	80
	<i>oloca</i>	<i>naca</i>	Decreased neutrophils	Sudan black B staining	Reduced survival of stromal cells	81

and facilitated the reliable and cost-effective identification of causative mutations from forward genetic screens. These ongoing developments and refinements in genetic mapping methods are driving advancements that are improving and reinforcing the knowledge of vertebrate embryogenesis and human disease.

### PRIMITIVE HEMATOPOIETIC GENES IDENTIFIED VIA FORWARD GENETICS

Over the past 30 years, forward genetic screening studies using zebrafish have identified numerous genes and elucidated the genetic mechanisms involved in hematopoiesis. This section provides examples of the studies that have elucidated the specific roles of genes identified by forward screening in distinct types of hematopoiesis during zebrafish embryogenesis (Table 1).

#### Erythropoiesis

Primitive hematopoiesis is regulated by a complex interplay of genes, and this area of research is constantly evolving. Nusslein-Volhard et al. conducted a pioneering forward genetic screening study in the 1990s to characterize the phenotypes of mutants associated with primitive hematopoiesis<sup>12,16,57</sup>. Subsequently,

mapping identified the relevant genes, and further research was conducted to determine the causes of hematopoietic defects in these mutants<sup>58,59</sup>. For instance, by confirming ENU-induced mutations based on GATA-1/GATA-2 expression and analyzing isolated blood cells, a *moonshine* (*mon*) mutant with a bloodless phenotype in the embryo was obtained<sup>12</sup>. Subsequent investigation identified the *mon* gene as transcriptional intermediary factor 1 gamma (*tif1γ*) through positional cloning using microsatellite markers, revealing that mutations in *tif1γ* trigger apoptosis in erythroid progenitor cells<sup>58</sup>.

In addition to *tif1γ*, the *gata1* gene has also been identified as a primitive erythropoiesis marker through forward genetic screening. The initial mutant, named *vlad tepes* (*vlt*), was isolated by visually observing a decrease in the number of red blood cells<sup>16</sup>. Diploid mapping using centromeric markers confirmed that *vlt* is the *gata1* gene<sup>59</sup>. Further investigation revealed that the mutated form of *gata1* could not bind to the promoter of its target genes, resulting in erythropoietic defects. Another erythropoietic gene identified through forward screening is *tal1*, which was mapped using SSLP markers<sup>60,61</sup>. Initially discovered in a screening study to identify genes involved in blood vessel development, subsequent studies revealed that *tal1* plays a critical role in red blood cell

differentiation, with a lack of *tal1* leading to anemia<sup>60,62</sup>. This discovery implies that hemangioblasts, which are precursors of primitive hematopoiesis, can differentiate into endothelial cells. Consequently, mutations affecting blood vessel development can also impact primitive hematopoiesis<sup>3,23</sup>.

#### Myelopoiesis

Forward genetic screening has also yielded mutants that cause defects in primitive myelopoiesis, which is crucial for early embryonic immunity<sup>23,27</sup>. Liu et al. conducted ENU mutagenesis and screened mutant embryos via WISH using probes for the myeloid cell markers *cbfb* and *l-plastin*<sup>63,64</sup>. A mutant named *mummy* (*mmy*) with reduced expression of both myeloid markers was identified. The *mmy* mutant displayed no blood circulation and a reduction in markers related to erythropoiesis, indicating that both myelopoiesis and erythropoiesis were affected. Subsequent positional cloning of the *mmy* mutant identified the causative mutation in the splicing factor gene *dhx8*<sup>64</sup>.

Another mutant, *moli*, related to primitive myelopoiesis, was generated by ENU mutagenesis and screened using neutral red and Sudan Black to identify macrophages and neutrophils, respectively<sup>57</sup>. These mutants showed a reduction in both macrophages and neutrophils and induction of cell cycle arrest in myeloid progenitor cells. Through linkage mapping using SSLP markers, the gene encoded by the *moli* mutant gene was found to be a bZIP transcription factor gene, *cebpa*, indicating that Cebpa plays an essential role in the maintenance of neutrophils during zebrafish embryogenesis.

### FORWARD SCREENING FOR HSC DEVELOPMENT

Given that blood cell lineages are generated through HSC development, understanding the detailed molecular mechanisms of HSC emergence and differentiation is crucial. Numerous key genes associated with HSC development have been identified through forward genetic approaches involving ENU-induced mutations. For instance, the functional role of *cmyb*, a widely recognized HSC marker, was identified through ENU-induced forward genetic screening<sup>65,66</sup>. Zebrafish embryos carrying *cmyb* mutations show failure of definitive hematopoiesis, indicating the essential role of this gene in regulating HSC development. Additionally, various forward screening studies have provided valuable insights into the function of novel genes related to HSC precursor formation, HSC specification and maintenance, and the regulation of HSC development by environmental signals.

#### Early endothelial and blood cells

Many genetic aberrations associated with HSC emergence and differentiation have been identified through ENU mutagenesis screening. For example, the *tal1* mutant phenotype exhibited a profound underdevelopment of the vasculature with the loss of not only *gata1*, a primitive erythrocyte marker, but also *runx1*, an early definitive HSC marker<sup>61</sup>. In addition, the *cloche* mutant embryos displayed cardiac abnormalities with a substantial reduction in both blood and vasculature formation<sup>35</sup>. The absence of GATA-2 expression, a key marker for early hematopoietic stem and progenitor cells in intermediate cellular mass (ICM), underscores the indispensable role of *cloche* in HSC generation<sup>26</sup>. Recent studies have revealed that the causative gene for *cloche*, which plays a significant role in both vascular and blood specification, is a PAS domain-containing bHLH transcription factor<sup>67</sup>.

Other mutants, such as *plcy1*, *tbx16*, and *hdac1*, manifest defective axial vascular organization, artery specification, and vessel or artery formation, respectively<sup>68</sup>. Notably, all of these mutants exhibit a deficiency in HSC formation, as evidenced by the complete loss of *runx1* expression. These genetic anomalies collectively accentuate the interplay between vascular and hematopoietic morphogenesis. Therefore, the development of the vasculature, including the DA, veins, and arteries, emerges as a fundamental orchestrator of HSC generation from endothelial cells

in the ventral wall of the DA. Consequently, ensuring proper vascular morphogenesis is necessary for the initiation and maintenance of HSCs.

#### HSC specification

Forward genetic approaches have identified genes that specifically impact the formation and specification of definitive HSCs without affecting primitive hematopoiesis and vasculogenesis. Recently, we identified a novel gene, *supt16h*, using forward genetics<sup>69</sup>. The *supt16h* mutants displayed an enlarged hindbrain and a deformed tail at 32 hpf. These traits were linked to the failure of HSC formation and proliferation, and the mutants displayed normal formation of upstream or adjacent tissues and cells, including posterior lateral mesoderm, sclerotome, vasculature, primitive erythrocytes, and primitive leukocytes. Furthermore, the loss of *supt16h* disrupts Notch activation, a pathway that is crucial for HSC emergence and specification. Subsequent investigations revealed that the absence of *supt16h* leads to increased expression of *p53*, which elevates the expression of *phc1*, a Notch repressor gene, resulting in the inhibition of Notch activation and HSC formation. This study demonstrated the significant role of *supt16h* in the Notch pathway and the regulation of HSC development. The study also exemplifies how forward genetic screens can not only identify novel genes of interest but also guide further research on the molecular mechanisms relevant to hematopoietic development.

#### Expansion and maintenance of HSCs

Through forward genetic screening, numerous novel genes involved in the expansion and maintenance of definitive HSCs have been identified. Three distinct mutants generated by ENU mutagenesis (*rumba*, *samba*, and *tango*) were identified through reduced expression of thymic *rag1*, a definitive hematopoiesis marker. Subsequent positional cloning revealed that *rumba* encodes a novel zinc finger protein, *samba* encodes the mitotic spindle protein Haus3, and *tango* encodes SUMO-activating enzyme subunit 1 (*sae1*)<sup>70,71</sup>. Loss of *rumba* and *haus3* results in defective HSC maintenance due to alterations in the cell cycle profile, whereas loss of *sae1* reduces HSC survival owing to SUMOylation defects. Another ENU screening study using thymic *rag1* expression identified mutated RNA polymerase III component 9 (*rpc9*) in the mutant line with definitive hematopoiesis defects. Loss of *rpc9* leads to a defect in HSC maintenance by inducing p53-dependent apoptosis<sup>72</sup>. The *ceylon* (*cey*) mutants identified through definitive *rag1* WISH exhibit severe defects in HSC maintenance<sup>40,73</sup>. SSLP-based linkage mapping identified the causative mutation in the *cey* mutants within the *tbl3* gene, which is associated with cell cycle regulation. Moreover, *grechetto* mutants were identified by employing the pan-leukocyte marker *mpx* in the CHT at 5 dpf. Positional cloning revealed that cleavage and polyadenylation specificity factor 1 (*cpsf1*) is the mutated gene in the *grechetto* mutant<sup>74</sup>. Further characterization of the mutant suggested that proper mRNA maturation is needed for HSC survival during definitive hematopoiesis.

In addition, the *cmyb* WISH screening method was used to identify multiple genes related to HSC maintenance and expansion. Pan et al. obtained HSC-defective mutants exhibiting reduced *cmyb* expression via ENU mutagenesis and subsequently mapped the mutated genes using bulk segregation analysis along with SSLP markers. Three genes, *topbp1*, *kri1l*, and *gemin5*, were found to be linked to the mutant phenotypes<sup>75-77</sup>. Further characterization of these mutants showed that TopBP1 is involved in HSC maintenance through regulation of p53-dependent apoptosis in the CHT. Loss of the rRNA maturation factor *kri1l* leads to misfolded proteins in HSCs and induces autophagy via PERK signaling. *gemin5* mutants exhibited a significantly decreased number of HSCs in the CHT due to defective HSC proliferation. Additionally, in a separate study, *slc20a1b* was found to be essential for HSC expansion in the CHT through  $\beta$ e1-globin WISH forward screening and subsequent analysis of *cmyb* expression in the CHT<sup>78,79</sup>.

### Niche for HSC expansion

Forward genetic studies have provided valuable insights into the crucial roles of niches where HSC expansion and maintenance occur. Utilizing *cmyb* WISH screening, a genetic mutation in integrin alpha 4 (*itga4*) that causes severe defects in definitive hematopoiesis has been identified<sup>80</sup>. Further investigation using photoconvertible transgenic lines revealed that the HSC retention time in the CHT is reduced in the *itga4* mutants. The loss of *vcam1*, a ligand of the Itga4-Itgb1 complex, leads to similar defects in definitive hematopoiesis as those observed in *itga4* mutants. Notably, macrophages expressing *vcam1* emerge as critical players in HSC retention within the CHT niche, highlighting the importance of the niche microenvironment in regulating HSC function<sup>80</sup>.

In another forward genetic study, researchers identified the *oloca* mutant using Sudan Black staining, which exhibited a decreased number of neutrophils compared to wild-type embryos at 3.5 dpf<sup>81</sup>. The *oloca* mutant carries a mutation in the *naca* gene, resulting in a reduced number of HSCs in the CHT, accompanied by abnormal proliferation and differentiation. Further investigations showed that *Naca* within the stromal niche plays a non-cell autonomous role in facilitating the settlement and formation of HSCs in the CHT. These findings highlight the significant genetic factors influencing HSC expansion and underscore the pivotal role of the CHT niche in regulating HSC function.

### SUMMARY AND CONCLUSION

In recent years, zebrafish have emerged as a powerful animal model for studying the genetic basis of hematopoiesis. Various gene editing and analysis techniques, including forward genetic screening, reverse screening, CRISPR-based approaches, and NGS, have been employed to identify genes associated with hematopoiesis and blood disorders. Notably, forward genetic screening, exemplified by methods such as ENU mutagenesis combined with linkage mapping techniques, has proven particularly advantageous in the zebrafish model. This strategy has facilitated the unbiased identification of numerous genes and pathways that regulate various aspects of blood development. In this review, we have explored the diverse genes and pathways involved in hematopoietic processes, providing critical insights into the molecular mechanisms underlying hematopoiesis via forward genetic screening.

As mutant mapping technologies advance and our knowledge of zebrafish genetics expands, we anticipate that this genetic model system will continue to revolutionize our understanding of hematopoiesis and blood disorders. Recent analyses using NGS and genome-wide association studies (GWASs) on extensive databases have highlighted potential blood disease-related genes and their associated pathological mechanisms<sup>82,83</sup>. Building on data-driven studies, zebrafish gene editing via the CRISPR system enhances our understanding of the related disease mechanisms and facilitates the development of disease models<sup>84</sup>. Collectively, these multifaceted gene editing approaches epitomize the utility of the zebrafish model in deciphering the complexities of hematopoiesis and its relevance to human health<sup>3</sup>.

### REFERENCES

- Bahary, N. & Zon, L. I. Use of the zebrafish (*Danio rerio*) to define hematopoiesis. *Stem Cells* **16 Suppl 2**, 67–78 (1998).
- Traver, D. et al. Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat. Immunol.* **4**, 1238–1246 (2003).
- Gore, A. V., Pillay, L. M., Venero Galanternik, M. & Weinstein, B. M. The zebrafish: a finstactic model for hematopoietic development and disease. *Wiley Interdiscip. Rev. Dev. Biol.* **7**, e312 (2018).
- Nasevicius, A. & Ekker, S. C. Effective targeted gene 'knockdown' in zebrafish. *Nat. Genet.* **26**, 216–220 (2000).
- Bedell, V. M., Westcot, S. E. & Ekker, S. C. Lessons from morpholino-based screening in zebrafish. *Brief. Funct. Genom.* **10**, 181–188 (2011).

- Auer, T. O. & Del Bene, F. CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. *Methods* **69**, 142–150 (2014).
- Shah, A. N., Davey, C. F., Whitebirch, A. C., Miller, A. C. & Moens, C. B. Rapid reverse genetic screening using CRISPR in zebrafish. *Nat. Methods* **12**, 535–540 (2015).
- Blackburn, P. R., Campbell, J. M., Clark, K. J. & Ekker, S. C. The CRISPR system—keeping zebrafish gene targeting fresh. *Zebrafish* **10**, 116–118 (2013).
- Li, M., Zhao, L., Page-McCaw, P. S. & Chen, W. Zebrafish genome engineering using the CRISPR-Cas9 system. *Trends Genet.* **32**, 815–827 (2016).
- Petri, K. et al. CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nat. Biotechnol.* **40**, 189–193 (2022).
- Rosello, M., Serafini, M., Concordet, J. P. & Del Bene, F. Precise mutagenesis in zebrafish using cytosine base editors. *Nat. Protoc.* **18**, 2794–2813 (2023).
- Ransom, D. G. et al. Characterization of zebrafish mutants with defects in embryonic hematopoiesis. *Development* **123**, 311–319 (1996).
- Driever, W. et al. A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123**, 37–46 (1996).
- Haffter, P. et al. The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **123**, 1–36 (1996).
- Voz, M. L. et al. Fast homozygosity mapping and identification of a zebrafish ENU-induced mutation by whole-genome sequencing. *PLoS One* **7**, e34671 (2012).
- Weinstein, B. M. et al. Hematopoietic mutations in the zebrafish. *Development* **123**, 303–309 (1996).
- Patton, E. E. & Zon, L. I. The art and design of genetic screens: zebrafish. *Nat. Rev. Genet.* **2**, 956–966 (2001).
- Raby, L., Volkel, P., Le Bourhis, X. & Angrand, P. O. Genetic Engineering of Zebrafish in Cancer Research. *Cancers (Basel)* **12**, <https://doi.org/10.3390/cancers12082168> (2020).
- Moens, C. B., Donn, T. M., Wolf-Saxon, E. R. & Ma, T. P. Reverse genetics in zebrafish by TILLING. *Brief. Funct. Genom. Proteom.* **7**, 454–459 (2008).
- Henke, K., Bowen, M. E. & Harris, M. P. Perspectives for identification of mutations in the zebrafish: making use of next-generation sequencing technologies for forward genetic approaches. *Methods* **62**, 185–196 (2013).
- Davey, J. W. et al. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* **12**, 499–510 (2011).
- de Jong, J. L. & Zon, L. I. Use of the zebrafish system to study primitive and definitive hematopoiesis. *Annu Rev. Genet.* **39**, 481–501 (2005).
- Chen, A. T. & Zon, L. I. Zebrafish blood stem cells. *J. Cell Biochem* **108**, 35–42 (2009).
- Bertrand, J. Y., Cisson, J. L., Stachura, D. L. & Traver, D. Notch signaling distinguishes 2 waves of definitive hematopoiesis in the zebrafish embryo. *Blood* **115**, 2777–2783 (2010).
- Hadland, B. K. et al. A requirement for Notch1 distinguishes 2 phases of definitive hematopoiesis during development. *Blood* **104**, 3097–3105 (2004).
- Davidson, A. J. & Zon, L. I. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. *Oncogene* **23**, 7233–7246 (2004).
- Forrester, A. M., Berman, J. N. & Payne, E. M. Myelopoiesis and myeloid leukemogenesis in the zebrafish. *Adv. Hematol.* **2012**, 358518 (2012).
- Thompson, M. A. et al. The cloche and spadetail genes differentially affect hematopoiesis and vasculogenesis. *Dev. Biol.* **197**, 248–269 (1998).
- Burns, C. E. et al. Isolation and characterization of runxa and runxb, zebrafish members of the runt family of transcriptional regulators. *Exp. Hematol.* **30**, 1381–1389 (2002).
- Bertrand, J. Y. et al. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* **464**, 108–111 (2010).
- Kissa, K. & Herbomel, P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature* **464**, 112–115 (2010).
- Murayama, E. et al. Tracing hematopoietic precursor migration to successive hematopoietic organs during zebrafish development. *Immunity* **25**, 963–975 (2006).
- Jin, H., Xu, J. & Wen, Z. Migratory path of definitive hematopoietic stem/progenitor cells during zebrafish development. *Blood* **109**, 5208–5214 (2007).
- Kissa, K. et al. Live imaging of emerging hematopoietic stem cells and early thymus colonization. *Blood* **111**, 1147–1156 (2008).
- Stainier, D. Y. Zebrafish genetics and vertebrate heart formation. *Nat. Rev. Genet.* **2**, 39–48 (2001).
- Solnica-Krezel, L., Schier, A. F. & Driever, W. Efficient recovery of ENU-induced mutations from the zebrafish germline. *Genetics* **136**, 1401–1420 (1994).
- Knapik, E. W. ENU mutagenesis in zebrafish—from genes to complex diseases. *Mamm. Genome* **11**, 511–519 (2000).
- Solnica-Krezel, L. et al. Mutations affecting cell fates and cellular rearrangements during gastrulation in zebrafish. *Development* **123**, 67–80 (1996).
- Thisse, C. & Thisse, B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3**, 59–69 (2008).
- Trede, N. S. et al. Zebrafish mutants with disrupted early T-cell and thymus development identified in early pressure screen. *Dev. Dyn.* **237**, 2575–2584 (2008).

41. Lawson, N. D. & Wolfe, S. A. Forward and reverse genetic approaches for the analysis of vertebrate development in the zebrafish. *Dev. Cell* **21**, 48–64 (2011).
42. Clements, W. K. & Traver, D. Signalling pathways that control vertebrate haematopoietic stem cell specification. *Nat. Rev. Immunol.* **13**, 336–348 (2013).
43. Jin, S. W. et al. A transgene-assisted genetic screen identifies essential regulators of vascular development in vertebrate embryos. *Dev. Biol.* **307**, 29–42 (2007).
44. Chi, N. C. et al. Genetic and physiologic dissection of the vertebrate cardiac conduction system. *PLoS Biol.* **6**, e109 (2008).
45. Lawson, N. D. & Weinstein, B. M. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev. Biol.* **248**, 307–318 (2002).
46. Covassin, L. D. et al. A genetic screen for vascular mutants in zebrafish reveals dynamic roles for Vegf/Plcg1 signaling during artery development. *Dev. Biol.* **329**, 212–226 (2009).
47. Leshchiner, I. et al. Mutation mapping and identification by whole-genome sequencing. *Genome Res.* **22**, 1541–1548 (2012).
48. Obholzer, N. et al. Rapid positional cloning of zebrafish mutations by linkage and homozygosity mapping using whole-genome sequencing. *Development* **139**, 4280–4290 (2012).
49. Sarin, S., Prabhhu, S., O'Meara, M. M., Pe'er, I. & Hobert, O. Caenorhabditis elegans mutant allele identification by whole-genome sequencing. *Nat. Methods* **5**, 865–867 (2008).
50. Blumenstiel, J. P. et al. Identification of EMS-induced mutations in Drosophila melanogaster by whole-genome sequencing. *Genetics* **182**, 25–32 (2009).
51. Bowen, M. E., Henke, K., Siegfried, K. R., Warman, M. L. & Harris, M. P. Efficient mapping and cloning of mutations in zebrafish by low-coverage whole-genome sequencing. *Genetics* **190**, 1017–1024 (2012).
52. Miller, A. C., Obholzer, N. D., Shah, A. N., Megason, S. G. & Moens, C. B. RNA-seq-based mapping and candidate identification of mutations from forward genetic screens. *Genome Res.* **23**, 679–686 (2013).
53. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
54. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
55. Flicek, P. et al. Ensembl 2011. *Nucleic Acids Res.* **39**, D800–806, (2011).
56. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
57. Dai, Y. et al. Cebpalpha is essential for the embryonic myeloid progenitor and neutrophil maintenance in zebrafish. *J. Genet. Genom.* **43**, 593–600 (2016).
58. Ransom, D. G. et al. The zebrafish moonshine gene encodes transcriptional intermediary factor 1gamma, an essential regulator of hematopoiesis. *PLoS Biol.* **2**, E237 (2004).
59. Lyons, S. E. et al. A nonsense mutation in zebrafish gata1 causes the bloodless phenotype in vlad tepes. *Proc. Natl Acad. Sci. USA* **99**, 5454–5459 (2002).
60. Habeck, H. et al. Analysis of a zebrafish VEGF receptor mutant reveals specific disruption of angiogenesis. *Curr. Biol.* **12**, 1405–1412 (2002).
61. Bussmann, J., Bakkers, J. & Schulte-Merker, S. Early endocardial morphogenesis requires Scf/Tal1. *PLoS Genet* **3**, e140 (2007).
62. Qian, F. et al. Distinct functions for different scl isoforms in zebrafish primitive and definitive hematopoiesis. *PLoS Biol.* **5**, e132 (2007).
63. Blake, T. et al. Zebrafish homolog of the leukemia gene CBFb: its expression during embryogenesis and its relationship to scl and gata-1 in hematopoiesis. *Blood* **96**, 4178–4184 (2000).
64. English, M. A. et al. Incomplete splicing, cell division defects, and hematopoietic blockage in dhx8 mutant zebrafish. *Dev. Dyn.* **241**, 879–889 (2012).
65. Paik, E. J. & Zon, L. I. Hematopoietic development in the zebrafish. *Int J. Dev. Biol.* **54**, 1127–1137 (2010).
66. Soza-Ried, C., Hess, I., Netuschil, N., Schorpp, M. & Boehm, T. Essential role of c-myc in definitive hematopoiesis is evolutionarily conserved. *Proc. Natl Acad. Sci. USA* **107**, 17304–17308 (2010).
67. Reischauer, S. et al. Cloche is a bHLH-PAS transcription factor that drives haemato-vascular specification. *Nature* **535**, 294–298 (2016).
68. Burns, C. E. et al. A genetic screen in zebrafish defines a hierarchical network of pathways required for hematopoietic stem cell emergence. *Blood* **113**, 5776–5782 (2009).
69. Espanola, S. G. et al. Haematopoietic stem cell-dependent Notch transcription is mediated by p53 through the Histone chaperone Supt16h. *Nat. Cell Biol.* **22**, 1411–1422 (2020).
70. Du, L. et al. Rumba and Haus3 are essential factors for the maintenance of hematopoietic stem/progenitor cells during zebrafish hematopoiesis. *Development* **138**, 619–629 (2011).
71. Li, X., Lan, Y., Xu, J., Zhang, W. & Wen, Z. SUMO1-activating enzyme subunit 1 is essential for the survival of hematopoietic stem/progenitor cells in zebrafish. *Development* **139**, 4321–4329 (2012).
72. Wei, Y., Xu, J., Zhang, W., Wen, Z. & Liu, F. RNA polymerase III component Rpc9 regulates hematopoietic stem and progenitor cell maintenance in zebrafish. *Development* **143**, 2103–2110 (2016).
73. Hutchinson, S. A. et al. Tbl3 regulates cell cycle length during zebrafish development. *Dev. Biol.* **368**, 261–272 (2012).
74. Bolli, N. et al. cpsf1 is required for definitive HSC survival in zebrafish. *Blood* **117**, 3996–4007 (2011).
75. Gao, L. et al. TopBP1 governs hematopoietic stem/progenitor cells survival in zebrafish definitive hematopoiesis. *PLoS Genet* **11**, e1005346 (2015).
76. Jia, X. E. et al. Mutation of kri11 causes definitive hematopoiesis failure via PERK-dependent excessive autophagy induction. *Cell Res.* **25**, 946–962 (2015).
77. Liu, X. et al. Mutation of Gemin5 causes defective hematopoietic stem/progenitor cells proliferation in zebrafish embryonic hematopoiesis. *Front Cell Dev. Biol.* **9**, 670654 (2021).
78. Chen, J. et al. Slc20a1b is essential for hematopoietic stem/progenitor cell expansion in zebrafish. *Sci. China Life Sci.* **64**, 2186–2201 (2021).
79. Wang, K. et al. Large-scale forward genetic screening analysis of development of hematopoiesis in zebrafish. *J. Genet. Genom.* **39**, 473–480 (2012).
80. Li, D. et al. VCAM-1(+) macrophages guide the homing of HSPCs to a vascular niche. *Nature* **564**, 119–124 (2018).
81. Murayama, E. et al. NACA deficiency reveals the crucial role of somite-derived stromal cells in haematopoietic niche formation. *Nat. Commun.* **6**, 8375 (2015).
82. Manier, S., Salem, K., Glavey, S. V., Roccaro, A. M. & Ghobrial, I. M. Genomic aberrations in multiple myeloma. *Cancer Treat. Res.* **169**, 23–34 (2016).
83. Boddu, P. C. & Kadia, T. M. Molecular pathogenesis of acquired aplastic anemia. *Eur. J. Haematol.* **102**, 103–110 (2019).
84. Xie, Y. et al. ARHGEF12 regulates erythropoiesis and is involved in erythroid regeneration after chemotherapy in acute lymphoblastic leukemia patients. *Haematologica* **105**, 925–936 (2020).

## AUTHOR CONTRIBUTIONS

H.S., U.S., U.N., and Y.L. discussed and commented on the paper. H.S. and Y.L. created the figures. H.S. and Y.L. revised and reviewed the paper.

## FUNDING INFORMATION

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (2022R1A2C100677812). The illustrations were generated using BioRender.com.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Yoonsung Lee.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.