# COMPOSITIONS AND METHODS FOR MED26-MEDIATED REGULATION OF ERYTHROPOIESIS

## CROSS-REFERENCE TO RELATED APPLICATIONS

**[0001]** This application claims priority to International Application No. PCT/CN2023/100427 filed on June 15, 2023, the disclosure of which is incorporated herein by reference in its entirety.

## FIELD OF THE INVENTION

**[0002]** The present application is in the field of transcriptional gene regulation.

## BACKGROUND

**[0003]** In steady-state erythropoiesis, approximately 2-3 million red blood cells are produced per second in the bone marrow(4). The early erythroid progenitors, burst-forming unit-erythroid (BFU-E) cells, can differentiate into the late progenitors, colony-forming unit-erythroid (CFU-E) cells(5). CFU-Es progressively differentiate into mature red blood cells through a series of 3–5 cell divisions, termed terminal erythropoiesis(6). During terminal erythropoiesis, erythroblasts undergo dramatic changes including nuclear condensation, wide-range transcription repression, extensive hemoglobin biogenesis, enucleation, and organelle clearance(7, 8). Defects in erythroid differentiation have been observed in myelodysplastic syndromes and megaloblastic anemia (9). Terminal erythropoiesis has been associated with a decline in certain histone marks involved with transcription elongation, including H3K36me2, H3K36me3, and H3K79me2 (1, 10), along with an increase in H4K20me, a modification associated with RNA polymerase II (Pol II) pausing and erythroblast chromatin condensation (11, 12). However, this process does not increase repressive histone marks H3K27me3 and H3K9me3, which are associated with heterochromatin formation. The positive transcription elongation factor (P-TEFb), which contains the catalytic subunit cyclin-dependent kinase 9 (CDK9), cooperates with the master erythroid transcription factor GATA1 to enhance transcription elongation (13). Hexim1, a regulator that promotes Pol II pausing, is highly expressed during terminal erythropoiesis, and is associated with accelerated differentiation of hexamethylene bisacetamide (HMBA)-treated HUDEP-2 erythroid cells(14). Collectively, these findings suggest regulatory roles in erythropoiesis related to transcription pausing and elongation.

**[0004]** Phase separation, which mediates membraneless compartment organization termed biomolecular condensates, has become an emerging model to explain diverse cellular events, including transcription regulation(15-17). Phase separation participates in transcription initiation and elongation via phosphorylation of the Pol II CTD(18). The presence of promoter condensates and gene-body condensates have been proposed at different transcription steps (3, 19). Although several proteins with phase separation capacity, including BRD4, MED1, and Pol II CTD(20-22), have been identified in transcription condensates, it remains unclear whether the dynamic composition of these condensates could contribute to driving the progression of developmental processes.

**[0005]** The Mediator complex, also known as the TRAP/SMCC, CRSP, PC2, or ARC complex, is a large multi-subunit complex composed of head, middle, tail, and CDK8 kinase modules, and is conserved from yeasts to metazoans (23),(24). The Mediator complex forms a functional bridge between gene promoters and enhancers, linking tissue-specific transcription factors (TFs) with general transcription factors (GTFs) and Pol II, thereby serving as an integrative hub for pre-initiation complex assembly, transcription elongation, and termination (23, 25). Several Mediator subunits have been shown to be important for various developmental processes through their associations with tissue-specific TFs(26). During erythropoiesis, MED1 is a cofactor of GATA1; MED1 knockout mice died at E11.5 of severe anemia (27, 28). Previous studies on Mediator-regulated developmental processes often focused on the function of a single subunit and its cooperation with a TF; however, it remains unclear whether individual Mediator subunits have differential functions throughout the stages of development, thereby contributing to the establishment of a context-dependent transcriptional program(29).

**[0006]** MED26 is a unique subunit in that, based on previous biochemical studies, its presence is usually exclusive to the CDK8 kinase module, and therefore is often regarded as a transcription activator (30). MED26 directly interacts with the super elongation complex (SEC) and the little elongation complex (LEC) containing P-TEFb via its N-terminal domain (NTD) (31-33). MED26 also functions as a molecular switch from its initiation state to its elongation state via interaction with the GTF TFIID (32). However, further investigation is necessary to determine the underlying molecular mechanism between MED26 and various aspects of transcription and its subsequent connection to the developmental process.

## SUMMARY

**[0007]** Provided is a method of promoting erythroid differentiation, comprising administering to an erythroid precursor cell an agent that increases RNA polymerase II pausing mediated by a mediator complex subunit (MED26) polypeptide in the cell.

**[0008]** In certain embodiments, the agent increases the interaction of the MED26 with a transcription pausing factor, such as a factor in a negative elongation factor (NELF), DRB Sensitivity Inducing Factor (DSIF) or polymerase associated factor (PAF) complex, particularly a PAF1, in the cell.

**[0009]** In certain embodiments, the agent increases occupancy of a super-enhancer site by the MED26, such as a super-enhancer selected from CDK6, BCL2, HMGA1, MYB, RPS14, RPL36AL, and RPL27.

**[0010]** In certain embodiments, the agent comprises the MED26 polypeptide or an active fragment thereof, or a nucleic acid encoding the active fragment.

**[0011]** In certain embodiments, the active fragment comprises an intrinsically disordered region (IDR) of the MED26 polypeptide, more particularly, the active fragment comprises a polypeptide that is at least 90%, such as at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the amino acid sequence of SEQ ID NO: 17.

**[0012]** In certain embodiments, the active fragment comprises a TFIIS domain and an intrinsically disordered region (IDR) of the MED26 polypeptide, more particularly, the active fragment comprises a polypeptide that is at least 90%, such as at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the amino acid sequence of SEQ ID NO: 2.

**[0013]** In certain embodiments, the agent comprises the MED26 polypeptide or the nucleic acid encoding the MED26 polypeptide, more particularly, the MED26 polypeptide comprises a polypeptide that is at least 90%, such as at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the amino acid sequence of SEQ ID NO: 1.

**[0014]** Also provided is a method of promoting erythroid differentiation, comprising administering to an erythroid precursor a polypeptide comprising the amino acid sequence of SEQ ID NO: 17 or a polynucleotide encoding the polypeptide.

**[0015]** In certain embodiments, the polypeptide comprises the amino acid sequence of SEQ ID NO: 2.

**[0016]** In certain embodiments, the polypeptide comprises the amino acid sequence of SEQ ID NO: 1.

**[0017]** In certain embodiments, the erythroid precursor is a hematopoietic stem and progenitor cell (HSPC), such as a CD34+ HSPC.

**[0018]** In certain embodiments, the erythroid precursor is an erythroblast, such as an erythroblast derived from a CD34+ HSPC.

**[0019]** In certain embodiments, the agent is administered to the erythroid precursor in vitro.

**[0020]** In certain embodiments, the agent is administered to the erythroid precursor in vivo in a subject in need thereof.

**[0021]** In certain embodiments, the subject is in need of treating a disease associated with defective erythroid differentiation.

**[0022]** In certain embodiments, the disease associated with defective erythroid differentiation is a myelodysplastic syndrome, such as refractory anemia or refractory cytopenia, erythroid dysplasia, bone marrow failure or megaloblastic anemia.

**[0023]** Also provided is a method of identifying an agent for promoting erythroid differentiation, comprising:

a. providing a composition comprising a MED26 polypeptide or an active fragment thereof;
b. contacting the composition with a test agent; and
c. assessing whether the test agent increases the capacity of the MED26 to mediate RNA polymerase II pausing.

**[0024]** In certain embodiments, the assessing step comprises assessing whether the test agent promotes the capacity of the MED26 to form an aggregate, to interact with a transcription pausing factor, and/or to occupy a super-enhancer site.

**[0025]** Also provided is a composition comprising an agent that increases RNA polymerase II pausing mediated by a MED26 polypeptide.

**[0026]** Also provided is an agent that increases RNA polymerase II pausing mediated by a MED26 polypeptide for use in a method of promoting erythroid differentiation in an erythroid precursor cell.

**[0027]** Also provided is a method of delivering a reagent to a cell, comprising: (a) preparing a condensate comprising a MED26 polypeptide or active fragment thereof and the reagent; and (b)

contacting the cell with the condensate, whereby the MED26 polypeptide or active fragment thereof delivers the reagent to the cell.

**[0028]** In certain embodiments, the reagent is a nucleic acid, such as one or more RNA or DNA molecules.

**[0029]** In certain embodiments, the MED26 polypeptide or active fragment thereof comprises an amino acid sequence having at least 75%, such as at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% , 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%,, sequence identity to SEQ ID NO: 16. In certain embodiments, the MED26 polypeptide or active fragment thereof comprises an amino acid sequence having at least 75%, such as at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% , 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%, sequence identity to SEQ ID NOs: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15. In certain embodiments, the MED26 polypeptide or active fragment thereof comprises the amino acid sequence of SEQ ID NOs: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16. In certain embodiments, the MED26 polypeptide or active fragment thereof comprises the amino acid sequence of SEQ ID NO: 12. In certain embodiments, the MED26 polypeptide or active fragment thereof comprises the amino acid sequence of SEQ ID NO: 2.

**[0030]** In certain embodiments, the nucleic acid encodes one or more components for gene editing. In certain embodiments, the gene editing comprises a CRISPR gene editing, and the nucleic acid encodes one or more of a CRISPR RNA (crRNA), a tracrRNA that hybridizes with the crRNA, and a Cas endonuclease (such as Cas9, Cas12, Cas13), preferably, the nucleic acid encodes a single guide RNA comprising the crRNA and the tracrRNA and/or the Cas endonuclease, more preferably, the nucleic acid encodes the single guide RNA and a Cas9 endonuclease.

**[0031]** In certain embodiments, the nucleic acid is a DNA molecule, such as a plasmid DNA.

**[0032]** In certain embodiments, the nucleic acid is an RNA, such as an mRNA.

**[0033]** In certain embodiments, the reagent is a DNA molecule, such as a circular DNA (e.g., a plasmid DNA) or a linear DNA.

**[0034]** In certain embodiments, the reagent is an RNA, such as an mRNA, an siRNA, an antisense RNA, a linear RNA, a circular RNA or a tRNA.

**[0035]** In certain embodiments, the step of preparing the condensate comprises mixing the reagent with the MED26 polypeptide or active fragment thereof in a buffer comprising water.

**[0036]** In certain embodiments, the buffer further comprises polyethylene glycol (PEG).

**[0037]** In certain embodiments, the buffer comprises 1-30% (w/v) PEG, such as 2-20% (w/v) PEG, e.g., 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14% or 15% (w/v) PEG, preferably 10% (w/v) PEG.

**[0038]** In certain embodiments, the step of preparing the condensate comprises mixing the reagent with 1-200 μM, preferably 5-150 μM, such as 5, 10, 25, 50, 75, 100, 125, or 150 μM, more preferably 10 μM, of the MED26 polypeptide or active fragment thereof.

## BRIEF DESCRIPTION OF THE DRAWINGS

**[0039]** The foregoing and other objects, aspects, features, and advantages of exemplary embodiments will become more apparent and may be better understood by referring to the following description taken in conjunction with the accompanying drawings.

**[0040] FIGs. 1A-F** show transcription condensates switched to MED26 enriched form in erythroid differentiation. FIG. 1A shows Western blot analysis of relative abundance of several Mediator subunits over time (Days 0-20) in the human CD34+ erythroid culture with β-actin as a loading control. FIG. 1B shows Western blot analysis of relative abundance of BRD4 and Rpb1 over time (Days 0-20) in the human CD34+ erythroid culture with β-actin as a loading control. FIG. 1C shows on top a schematic diagram of the OptoDroplet assay and on bottom time-lapse images of various Mediator subunits analyzed by the OptoDroplet assay. The subunits are indicated on the left and subcellular localization is indicated on the right. FIG. 1D shows representative images of FRAP assay for EGFP-MED26 condensates in K562 cells. The normalized fluorescence intensity of EGFP-MED26 is represented as mean ± s.d. (n = 27, independent observations in 27 separated aggregates). FIG. 1E shows droplet formation assays of gradient diluted MED26 and MED1 IDR proteins under the low-salt buffer without PEG condition. FIG. 1F shows droplet formation assays in which 1,6-hexanediol (1,6-Hex) was added to test disruption of the liquid-like condensates. FIG. 1G shows droplet formation assays conducted in 25 mM Tris-HCl (pH 7.4) with various NaCl concentrations.

**[0041] FIGs. 2A-2I** show MED26 is essential to erythropoiesis under normal and PHZ-induced stress conditions. FIG. 2A shows a Western blot demonstrating MED26 knock-out efficiency in the mouse spleen cells with GAPDH as a loading control. FIG 2B shows images of the mice overview, femur and spleen of control and cko mice after MED26 knock-out induced by pIpC

injection. FIG. 2C shows Prussian blue staining for iron in control and cko mice spleen. FIG. 2D shows graphs of mean corpuscular volume (MCV) and mean corpuscular hemoglobin (MCH) in control and MED26 knock-out mice after 3 days pIpC injection induced knock-out. FIG. 2E shows Giemsa staining of peripheral blood smears from control and cko mice. Arrow indicates reticulocytes. FIG. 2F shows FACS analysis to detect differentiation states of mice femur bone marrow cells by two erythroid markers (CD71, Ter119). FIG. 2G shows levels of Hemoglobin (HGB), red blood cells (RBC), and hematocrit (HCT) in control and conditional MED26 knock-out (cKO) mice pretreated with 5 ug/g pIpC one day before PHZ (60 mg/kg) injection on day 0. FIG. 2H shows representative flow cytometry plots of Lin-, LSK, LT-HSC, ST-HSC, MPP, CMP, GMP, MEP cells in the bone marrow of control and cko mice, n=3. FIG. 2I shows percentages of Lin-, LSK, LT-HSC, ST-HSC, MPP, CMP, GMP, MEP cells in the bone marrow of control and cko mice. The p-values were calculated using an unpaired two-tailed Student's t-test (C,F,H), significant differences by Student's t-test are marked by asterisks: *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001, n.s., not significant.

**[0042] FIGs. 3A-3D** show MED26 is essential for the progression of normal erythropoiesis. FIG. 3A shows representative colony images for colony-forming assay of human CD34+ cells transduced with control shRNA or shRNA targeting MED26. FIG. 3B shows relative MED26 gene expression human CD34+ cells transduced with control shRNA or shRNA targeting MED26. FIGs. 3C and 3D show quantification of CFU-GM and CFU-GEMM colony size after knocking down MED26 in human CD34+ cells. Abbreviations: CFU-GEMM, colony-forming unit-granulocyte, erythrocyte, monocyte, megakaryocyte; CFU-GM, colony-forming unit-granulocyte, macrophage. The p-values were calculated using an unpaired two-tailed Student's t-test (D), significant differences by Student's t-test are marked by asterisks: *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001, n.s., not significant.

**[0043] FIGs. 4A-F** show phase-separation capacity of MED26 is associated with erythroid development. FIG. 4A shows a Western blot demonstrating MED26 over-expression efficiency in the ex vivo human CD34+ erythroid differentiation system on Day 8 with GAPDH as a loading control. FIG. 4B shows FACS analysis to detect two erythroid markers (CD71, CD235a) upon MED26 overexpression. FIG. 4C shows pictures of CD34+ cell pellets upon MED26 overexpression on Day 10 and Day 12. The dotted circles indicate the cell pellets position. FIG. 4D shows a schematic diagram of MED26 truncations of TFIIS domain, intrinsically disordered

region (IDR), and Mediator complex interaction domain. The results of in vivo condensate formation are summarized on the right. FIG. 4E shows representative time-lapse images of the indicated MED26 truncations in the OptoDroplet assay. FIG. 4F shows fluorescent imaging of an in vitro phase separation assay of the indicated MED26 truncations fused with EGFP. FIG. 4G shows FACS analysis to detect erythroid markers (CD71, CD235a) upon full-length or truncated MED26 overexpression in the ex vivo human CD34+ erythroid culture.

**[0044] FIGs. 5A-F** show RNA polymerase II exhibits transcription pausing at MED26-enriched loci. FIG. 5A shows heatmaps from CUT&Tag assays showing the number and distribution of MED1, MED26, GATA1 and GATA2 common peaks, MED1 unique peaks, and MED26 unique peaks on Day 4 in the primary human CD34+ derived erythroid culture. FIG. 5B shows the ratio of MED26 to MED1 signals on all the transcription start sites (TSS) with MED1 and/or MED26 occupancy in CUT&Tag assays of primary human erythroblasts on Day 4. Each dot represents one gene. FIG. 5C. shows a schematic defining the MED26-enriched or MED26-poor Mediator loci. FIG. 5D shows heatmaps showing RNA Pol II and PRO-seq signal of MED26-enriched or -poor genes from -3kb of TSS to +3kb of transcription end site (TES). FIG. 5E shows IGV visualization showing examples of MED26-enriched (top) and MED26-poor (bottom) loci. The grey shaded areas indicate MED26 occupancy around the TSS. NC indicates signals of negative control samples. RPB1 is the largest component of RNA Pol II. FIG. 5F shows boxplots comparing the Pausing Index of MED26-enriched or -poor loci. The pausing indices were calculated from Rpb1 CUT&Tag or PRO-seq. The p-values were calculated using the two-sided Wilcoxon rank-sum test.

**[0045] FIG. 6A-I** show MED26 enrichment promotes transcription pausing necessary for erythropoiesis. FIG. 6A shows a heatmap showing the relative interaction strength of indicated proteins baited by MED1 or MED26 in HEK293 cells. Relative interaction strength is calculated by NSAF from mass spectrometry data. FIG. 6B shows co-immunoprecipitation of HEK293 cells transfected with 3XFlag-MED1, 3XFlag-MED26, or 3XFlag-EV (control plasmid). Following immunopurification by Flag antibody, western blot was performed on the complexes using antibodies against components of the pausing complexes (NELF-A, NELF-D, PAF1, and LEO1), the Mediator CDK8 kinase module (CCNC1), and the elongation complex (CDK9 and CCNT1). FIG. 6C shows recombinant proteins of mCherry and mCherry-PAF1-400-531 (a.a. 400-531, the putative IDR of PAF1) were purified from the prokaryotic expression system and

detected by Coomassie blue staining. FIG. 6D shows *in vitro* phase separation assay of EGFP-fused MED26 1-480 and mCherry-PAF1-400-531, with a buffer containing 5% PEG-8000. FIG. 6E shows representative images of MED26 KO K562 cells transfected with MED26-truncations-EGFP and PAF1-mcherry. The white dotted circle indicates the nuclear area. FIG. 6F shows boxplots comparing the MED26 to MED1 signal ratio on Day 4 and Day 16 of the human CD34+ derived erythroid culture. (The p-values are calculated by paired Wilcox test). FIG. 6G shows boxplots comparing the Pausing Index on Day 4 and Day 16 of the human CD34+ derived erythroid culture. (The p-values are calculated by paired Wilcox test). FIG. 6H shows IGV visualization of the CUT&Tag signals of PRO-seq, MED1, and MED26 at RPS9 (non-erythroid gene) and HBB (erythroid gene) loci. FIG. 6I shows FACS analysis of erythroid markers CD71 and CD235a in primary human erythroblasts treated with DRB (5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole).

**[0046] FIG. 7** shows electron microscope images of RAW264.7 cell uptake of MED26-mRNA coacervates.

**[0047] FIG. 8** shows FACS analysis of Jurkat cells transfected with EGFP mRNA (Cy5-labeled) loaded MED26 coacervates formulations featuring varying PEG concentrations.

**[0048] FIG. 9A** shows fluorescence micrographs of 293T and Jurkat cells transfected with MED26-EGFP mRNA coacervates. Scale bar, 50 μm. FIG. 9B shows quantification of mRNA expression in 293T and Jurkat cells 24 hours post-transfection with MED26 coacervates or Lipofectamine MessegerMAX.

**[0049] FIG. 10A** shows schematic illustration of MED26 truncations. **FIG. 10B** shows quantification of mRNA expression in Jurkat cells 24 hours after transfection with coacervates formed by MED26 truncated proteins. **FIG. 10C** shows quantification of mRNA expression in Jurkat cells 24 hours after transfection with coacervates formulations featuring varying MED26 (135-480) concentrations. **FIG. 10D** shows quantification of mRNA expression in Jurkat cells 24 hours after transfection with coacervates formulations featuring varying higher MED26 (135-480) concentrations.

**[0050] FIG. 11** shows FACS analysis of Jurkat cells at 24 hours after treating with two mRNAs (EGFP mRNA and tdTomato mRNA) loaded coacervates.

**[0051] FIG. 12A** shows representative confocal laser scanning microscopy images of early endosome, coacervates (Cy5-labled mRNA), and nucleus (Hoechst 33342) in 293T cells 20

minutes post-transfection with MED26 coacervates. Scale bar, 5 μm. **FIG. 12B** shows representative confocal laser scanning microscopy images of early endosome , coacervates (Cy5-labled mRNA), and nucleus (Hoechst 33342) in 293T cells 1 hour post-transfection with MED26 coacervates. Scale bar, 5 μm. **FIG. 12C** shows temporal expression profile in Jurkat cells transfected with MED26-mRNA coacervates. **FIG. 12D** shows MED26 protein level over time in Jurkat cells transfected with MED26-mRNA coacervates.

**[0052] FIG. 13A** shows the relative quantity of mRNA levels in 293T-GFP cells 72 hours after treating with single-targeted siRNA-loaded MED26 coacervates. **FIG. 13B** shows relative quantity of mRNA levels in 293T-GFP cells 72 hours after treating with multiple-targeted siRNAs-loaded MED26 coacervates.

**[0053] FIG. 14A** shows quantification of plasmid (size: 8 kb and 15 kb) expression in Jurkat cells 48 hours post-transfection with MED26 coacervates or Lipofectamine 3000. **FIG. 14B** shows quantification of plasmid (size: 8 kb and 15 kb) expression in 293T cells 48 hours post-transfection with MED26 coacervates or Lipofectamine 3000. **FIG. 14C** shows quantification of plasmid expression in Jurkat cells 48 hours after transfection with coacervates packaging CRISPR/Cas9 pDNA with different MED26 (135-480) concentrations. **FIG. 14D** shows FACS analysis of Jurkat cells at 48 hours after treating with two plasmids (GFP plasmid and tdTomato plasmid)-packed coacervates. **FIG. 14E** shows a schematic illustration of the structure of the all-in-one pDNA-CRISPR/Cas9. **FIG. 14F** shows Transfection efficiency at the DNMT1, HBB and HPRT1 locus in Jurkat cells treated with single targeted pDNA-loaded MED26 coacervates. **FIG. 14G** shows indel frequencies at the DNMT1, HBB and HPRT1 locus in Jurkat cells treated with single targeted pDNA-loaded MED26 coacervates. **FIG. 14H** shows transfection efficiency at the DNMT1, HBB, HPRT1 locus in 293T cells treated with single or triple targeted pDNA-loaded MED26 coacervates. **FIG. 14I** shows indel frequencies at the DNMT1, HBB, HPRT1 locus in 293T cells treated with single or triple targeted pDNA-loaded MED26 coacervates.

## DETAILED DESCRIPTION

**[0054]** Various publications, articles and patents are cited or described in the background and throughout the specification; each of these references is herein incorporated by reference in its entirety. Discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is for the purpose of providing context for the invention.

Such discussion is not an admission that any or all of these matters form part of the prior art with respect to any inventions disclosed or claimed.

**[0055]** Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention pertains. Otherwise, certain terms used herein have the meanings as set forth in the specification.

**[0056]** It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise.

**[0057]** Unless otherwise stated, any numerical values, such as a concentration or a concentration range described herein, are to be understood as being modified in all instances by the term "about." Thus, a numerical value typically includes ± 10% of the recited value. For example, a concentration of 1 mg/mL includes 0.9 mg/mL to 1.1 mg/mL. Likewise, a concentration range of 1% to 10% (w/v) includes 0.9% (w/v) to 11% (w/v). As used herein, the use of a numerical range expressly includes all possible subranges, all individual numerical values within that range, including integers within such ranges and fractions of the values unless the context clearly indicates otherwise.

**[0058]** Unless otherwise indicated, the term "at least" preceding a series of elements is to be understood to refer to every element in the series. Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the invention.

**[0059]** As used herein, the terms "comprises," "comprising," "includes," "including," "has," "having," "contains" or "containing," or any other variation thereof, will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers and are intended to be non-exclusive or open-ended. For example, a composition, a mixture, a process, a method, an article, or an apparatus that comprises a list of elements is not necessarily limited to only those elements but can include other elements not expressly listed or inherent to such composition, mixture, process, method, article, or apparatus. Further, unless expressly stated to the contrary, "or" refers to an inclusive or and not to an exclusive or. For example, a condition A or B is satisfied by any one of the following: A is true (or present) and B is false (or not present), A is false (or not present) and B is true (or present), and both A and B are true (or present).

**[0060]** As used herein, the conjunctive term "and/or" between multiple recited elements is understood as encompassing both individual and combined options. For instance, where two elements are conjoined by "and/or," a first option refers to the applicability of the first element without the second. A second option refers to the applicability of the second element without the first. A third option refers to the applicability of the first and second elements together. Any one of these options is understood to fall within the meaning, and therefore satisfy the requirement of the term "and/or" as used herein. Concurrent applicability of more than one of the options is also understood to fall within the meaning, and therefore satisfy the requirement of the term "and/or."

**[0061]** As used herein, the term "consists of," or variations such as "consist of" or "consisting of," as used throughout the specification and claims, indicate the inclusion of any recited integer or group of integers, but that no additional integer or group of integers can be added to the specified method, structure, or composition.

**[0062]** As used herein, the term "consists essentially of," or variations such as "consist essentially of" or "consisting essentially of," as used throughout the specification and claims, indicate the inclusion of any recited integer or group of integers, and the optional inclusion of any recited integer or group of integers that do not materially change the basic or novel properties of the specified method, structure or composition. See M.P.E.P. § 2111.03.

**[0063]** As used herein, "subject" means any animal, preferably a mammal, most preferably a human. The term "mammal" as used herein, encompasses any mammal. Examples of mammals include, but are not limited to, cows, horses, sheep, pigs, cats, dogs, mice, rats, rabbits, guinea pigs, monkeys, humans, etc., more preferably a human.

**[0064]** The words "right," "left," "lower," and "upper" designate directions in the drawings to which reference is made.

**[0065]** It should also be understood that the terms "about," "approximately," "generally," "substantially," and like terms, used herein when referring to a dimension or characteristic of a component of the preferred invention, indicate that the described dimension/characteristic is not a strict boundary or parameter and does not exclude minor variations therefrom that are functionally the same or similar, as would be understood by one having ordinary skill in the art. At a minimum, such references that include a numerical parameter would include variations that, using mathematical and industrial principles accepted in the art (e.g., rounding, measurement or other systematic errors, manufacturing tolerances, etc.), would not vary the least significant digit.

**[0066]** The terms “identical” or percent “identity,” in the context of two or more nucleic acids or polypeptide sequences (e.g., MED26 protein and fragments thereof or polynucleotides that encode them), refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection.

**[0067]** For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

**[0068]** Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 1981; 2:482, by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 1970; 48:443, by the search for similarity method of Pearson & Lipman, Proc. Nat’l. Acad. Sci. USA 1988; 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by visual inspection (see generally, Current Protocols in Molecular Biology, F.M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., 1995 Supplement (Ausubel)).

**[0069]** Examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al., J. Mol. Biol. 1990; 215: 403-410 and Altschul et al., Nucleic Acids Res. 1997; 25: 3389-3402, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits act as seeds for initiating

searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased.

**[0070]** Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 1989; 89:10915).

**[0071]** In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 1993; 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

**[0072]** A further indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions.

**[0073]** As used herein, the term "isolated" means a biological component (such as a nucleic acid, peptide or protein) has been substantially separated, produced apart from, or purified away from

other biological components of the organism in which the component naturally occurs, i.e., other chromosomal and extrachromosomal DNA and RNA, and proteins. Nucleic acids, peptides and proteins that have been "isolated" thus include nucleic acids and proteins purified by standard purification methods. "Isolated" nucleic acids, peptides and proteins can be part of a composition and still be isolated if the composition is not part of the native environment of the nucleic acid, peptide, or protein. The term also embraces nucleic acids, peptides and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids.

**[0074]** As used herein, the term "polynucleotide," synonymously referred to as "nucleic acid molecule," "nucleotides" or "nucleic acids," refers to any polyribonucleotide or polydeoxyribonucleotide, which can be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that can be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short nucleic acid chains, often referred to as oligonucleotides.

**[0075]** As used herein, the term "vector" is a replicon in which another nucleic acid segment can be operably inserted so as to bring about the replication or expression of the segment.

**[0076]** As used herein, the term "host cell" refers to a cell comprising a nucleic acid molecule of the invention. The "host cell" can be any type of cell, e.g., a primary cell, a cell in culture, or a cell from a cell line. In one embodiment, a "host cell" is a cell transfected with a nucleic acid molecule of the invention. In another embodiment, a "host cell" is a progeny or potential progeny of such a transfected cell. A progeny of a cell may or may not be identical to the parent

cell, e.g., due to mutations or environmental influences that can occur in succeeding generations or integration of the nucleic acid molecule into the host cell genome.

**[0077]** The term "expression" as used herein, refers to the biosynthesis of a gene product. The term encompasses the transcription of a gene into RNA. The term also encompasses translation of RNA into one or more polypeptides, and further encompasses all naturally occurring post-transcriptional and post-translational modifications.

**[0078]** As used herein, the terms "peptide," "polypeptide," or "protein" can refer to a molecule comprised of amino acids and can be recognized as a protein by those of skill in the art. The conventional one-letter or three-letter code for amino acid residues is used herein. The terms "peptide," "polypeptide," and "protein" can be used interchangeably herein to refer to polymers of amino acids of any length. The polymer can be linear or branched, it can comprise modified amino acids, and it can be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), as well as other modifications known in the art.

**[0079]** The peptide sequences described herein are written according to the usual convention whereby the N-terminal region of the peptide is on the left and the C-terminal region is on the right. Although isomeric forms of the amino acids are known, it is the L-form of the amino acid that is represented unless otherwise expressly indicated.

**Methods of promoting erythroid differentiation**

**[0080]** Erythropoiesis is the process of making red blood cells. At each stage of maturation, erythroblasts are both phenotypically distinct and have a unique transcriptomic profile and chromatin landscape. The morphologic changes include a progressively acidophilic appearance due to accumulation of hemoglobin, a steady decrease in cell size, and dramatic nuclear condensation, which culminates in enucleation. This process requires the coordinated effort of epigenetic regulators and transcription factors, as well as precise regulation of RNA polymerase II activity. The average human makes two to three million red cells per second to maintain steady state, and avoid anemia. Defects in terminal erythroid maturation, such as nuclear

condensation defects, or asynchronous maturation of the nucleus and cytoplasm, are commonly found in myelodysplastic syndromes, and inherited anemias. Understanding the molecular mechanisms that govern terminal erythroid maturation is essential to understanding how mutations or other genetic perturbations result in dyserythropoiesis, and to designing rational therapies (Wells and Steiner, Front Genet. 2022; 13: 805265).

**[0081]** In a general aspect, the application provides methods of promoting erythroid differentiation.

**[0082]** In certain embodiments, a method of promoting erythroid differentiation comprises administering to an erythroid precursor cell an agent that increases RNA polymerase II pausing mediated by a mediator complex subunit 26 (MED26) polypeptide in the cell.

**[0083]** The term "stem cell" as used herein refers to undifferentiated cells of a multicellular organism having the ability to self-renew that can generate daughter cells that can undergo terminal differentiation into more than one distinct cell types with specific functions. A stem cell has the potential to differentiate into multiple types of cells and is capable of unlimited self-replication via asymmetric cell division, a process known as self-renewal. In preferred embodiment, the stem cell is an adult stem cell, also called a somatic stem cell, which is multipotent and can generate cell types within a specific lineage, such as blood cells or endothelial cells.

**[0084]** In certain embodiments, the stem cell is a hematopoietic stem cell. As used herein, the term “hematopoietic stem cells” (“HSCs”) refers to immature cells having the capacity to self-renew and to differentiate into one or more mature blood cells. Examples of mature blood cells include, but are not limited to, granulocytes (e.g., promyelocytes, neutrophils, eosinophils, basophils), erythrocytes (e.g., reticulocytes, erythrocytes), thrombocytes (e.g., megakaryoblasts, platelet producing megakaryocytes, platelets), monocytes, dendritic cells, microglia, osteoclasts, and lymphocytes.

**[0085]** The term “progenitor cell” as used herein refers to cells that are descendants of stem cells, which can further differentiate to create specialized cell types. The term is not limitative and does not limit these cells to a particular lineage. Unlike stem cells, progenitor cells have a lesser ability to self-renew. In addition, the cell potency of progenitor cells is usually more restricted than stem cells. A progenitor cell is normally only capable of differentiating into cells that belong to the same tissue or organ. Some progenitor cells have one final target cell that they

differentiate to (unipotent), while others have the potential to terminate in more than one cell type (oligopotent or multipotent).

**[0086]** As used herein, "an erythroid precursor cell" is any cell involved in the erythroid formation process. In certain embodiments, the erythroid precursor is a hematopoietic stem and progenitor cell (HSPC), such as a CD34+ HSPC. In certain embodiments, the erythroid precursor is an erythroblast, such as an erythroblast derived from a CD34+ HSPC.

**[0087]** It should be appreciated that the present invention contemplates the use of any agent that is capable of increasing RNA polymerase II pausing mediated by MED26 polypeptide. Any of a wide variety of agents can be used to increase RNA polymerase II pausing mediated by MED26 polypeptide. Without limitation, agents include small molecules, peptides, polypeptides, nucleic acids, oligonucleotides, antibodies, etc. In certain embodiments, the increased interaction or occupancy can be by increased affinity and/or increased amount or concentration.

**[0088]** As used herein, the term "MED26" or "MED26 polypeptide" refers to mediator of RNA polymerase II transcription subunit 26, or mediator complex subunit 26 protein, a component or subunit of the CRSP (cofactor required for SP1 activation) complex. Embodiments of the application are related to a MED26 polypeptide or an active fragment thereof. Preferably, the MED26 is a mammalian MED26, such as a human MED26. In certain embodiments, the MED26 polypeptide is a human MED26 polypeptide. In certain embodiments, the human MED26 polypeptide comprises the amino acid sequence of SEQ ID NO: 1 or an active fragment thereof.

**[0089]** In certain embodiments, the agent comprises the MED26 polypeptide or an active fragment thereof, or a nucleic acid encoding the active fragment. As used herein, the term "active fragment" refers to a fragment of MED26 or derivative thereof that is capable of interacting with a transcription pausing factor to mediate RNA polymerase II pausing or that is capable of forming a condensate with a reagent and delivering the reagent to a cell.

**[0090]** In certain embodiment, the active fragment comprises an intrinsically disordered region (IDR) of the MED26 polypeptide, more particularly, the active fragment comprises a polypeptide that is at least 75%, such as at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% , 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the amino acid sequence of SEQ ID NO: 17. In certain embodiments, the active fragment comprises a TFIIS domain and an intrinsically disordered region (IDR) of the MED26 polypeptide, more particularly, the active fragment comprises a polypeptide that is at least 75%, such as at least

75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% , 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the amino acid sequence of SEQ ID NO: 2.

**[0091]** In certain embodiments, the agent comprises the MED26 polypeptide or the nucleic acid encoding the MED26 polypeptide, more particularly, the MED26 polypeptide comprises a polypeptide that is at least 75%, such as at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% , 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the amino acid sequence of SEQ ID NO: 1.

**[0092]** A MED26 polypeptide or an active fragment thereof can be produced by any suitable methods in view of the present disclosure. In some embodiments, a MED26 polypeptide or an active fragment thereof is produced by recombinant production, e.g., encoded by a nucleic acid comprising the polynucleotide sequence of GenBank Accession No. NM_004831.5, or a fragment thereof. In certain embodiments, the nucleic acid can be optimized for protein expression.

**[0093]** As used herein, “enhancer” refers to a short region of DNA to which proteins (e.g., transcription factors) bind to enhance transcription of a gene. As used herein, "transcriptional coactivator" refers to a protein or complex of proteins that interacts with transcription factors to stimulate transcription of a gene. As used herein, a “super-enhancer” or “super-enhancer site” is a region of DNA comprising two or more enhancers that is collectively bound by an array of transcription factor proteins to drive transcription of genes involved in cell identity. Examples of super-enhancers include, but are not limited to, those described, e.g., in U.S. Patent Publication US2014/0287932, the content of which is incorporated herein by reference in its entirety .

**[0094]** In certain embodiments, the agent increases the interaction of the MED26 with a transcription pausing factor, such as a factor in a negative elongation factor (NELF), DRB Sensitivity Inducing Factor (DSIF) or polymerase associated factor (PAF) complex, particularly a PAF1, in the cell.

**[0095]** In certain embodiments, the agent increases occupancy of a super-enhancer site by the MED26, such as a super-enhancer for a gene selected from cyclin-dependent kinase 6 (CDK6), Apoptosis regulator Bcl-2 (BCL2), High mobility group protein HMG-I/HMG-Y (HMGA1), Transcriptional activator Myb (MYB), 40S ribosomal protein S14 (RPS14), 60S ribosomal protein L36a-like R (PL36AL), and 60S ribosomal protein L27 (RPL27).

**[0096]** Also provided is a method of promoting erythroid differentiation, comprising administering to an erythroid precursor a polypeptide comprising the amino acid sequence of SEQ ID NO: 17 or a polynucleotide encoding the polypeptide.

**[0097]** In certain embodiments, the polypeptide comprises the amino acid sequence of SEQ ID NO: 2. In certain, embodiments, the polypeptide comprises the amino acid sequence of SEQ ID NO: 1.

**[0098]** In certain embodiments, the agent is administered to the erythroid precursor *in vitro*. In certain embodiments, the agent is administered to the erythroid precursor *in vivo* in a subject in need thereof.

**[0099]** In certain embodiments, the subject is in need of treating a disease associated with defective erythroid differentiation. In certain embodiments, the disease associated with defective erythroid differentiation is a myelodysplastic syndrome, such as refractory anemia or refractory cytopenia, erythroid dysplasia, bone marrow failure or megaloblastic anemia.

**[00100]** Also provided is a method of delivering a reagent to a cell, comprising:

a. preparing a condensate comprising a MED26 polypeptide or active fragment thereof and the reagent; and
b. contacting the cell with the condensate,

whereby the MED26 polypeptide or active fragment thereof delivers the reagent to the cell.

**[00101]** Condensates have been explored as potential delivery systems for various substances due to their ability to encapsulate molecules within their liquid droplets.

**[00102]** In certain embodiments, the reagent is a small molecule such as protein, peptide, RNAs, DNAs or CRISPR components for gene editing. In certain embodiments, the condensate provides a protective environment that enhances the reagents stability, solubility, and/or controlled release. In certain embodiments, the reagent is a nucleic acid, such as one or more RNA or DNA molecules.

**[00103]** In certain embodiments, the nucleic acid encodes one or more components for gene editing. In certain embodiments, the gene editing comprises a CRISPR gene editing, and the nucleic acid encodes one or more of a CRISPR RNA (crRNA), a tracrRNA that hybridizes with the crRNA, and a Cas endonuclease (such as Cas9, Cas12, Cas13), preferably, the nucleic acid encodes a single guide RNA comprising the crRNA and the tracrRNA and/or the Cas

endonuclease, more preferably, the nucleic acid encodes the single guide RNA and a Cas9 endonuclease.

**[00104]** In certain embodiments, the nucleic acid is a DNA molecule, such as a plasmid DNA. the reagent is a DNA molecule, such as a circular DNA (e.g., a plasmid DNA) or a linear DNA (e.g., an antisense DNA).

**[00105]** In certain embodiments, the nucleic acid is an RNA, such as an mRNA. the reagent is an RNA, such as an mRNA, an siRNA, an antisense RNA, a linear RNA, a circular RNA or a tRNA. The RNA can be modified with one or more chemical modifications.

**[00106]** In certain embodiments, the MED26 polypeptide or active fragment thereof comprises an amino acid sequence having at least 75%, such as at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% , 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%, sequence identity to SEQ ID NOs: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16. In certain embodiments, the MED26 polypeptide or active fragment thereof comprises an amino acid sequence having at least 75%, such as at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% , 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%, sequence identity to SEQ ID NO: 16. In certain embodiments, the MED26 polypeptide or active fragment thereof comprises the amino acid sequence of SEQ ID NOs: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16. In certain embodiments, the MED26 polypeptide or active fragment thereof comprises the amino acid sequence of SEQ ID NO: 12. In certain embodiments, the MED26 polypeptide or active fragment thereof comprises the amino acid sequence of SEQ ID NO: 2.

**[00107]** In certain embodiments, the step of preparing the condensate comprises mixing the reagent with the MED26 polypeptide or active fragment thereof in a buffer comprising water.

**[00108]** In certain embodiments, the buffer further comprises polyethylene glycol (PEG). In certain embodiments, the buffer comprises 1-30% (w/v) PEG, such as 2-20% (w/v) PEG, e.g., 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14% or 15% (w/v) PEG, preferably 10% (w/v) PEG. In certain embodiments, the step of preparing the condensate comprises mixing the reagent with 1-200 μM, preferably 5-150 μM, such as 5, 10, 25, 50, 75, 100, 125, or 150 μM, more preferably 10 μM, of the MED26 polypeptide or active fragment thereof.

**Compositions**

**[00109]** Also provided herein are compositions (such as pharmaceutical compositions) comprising an agent that increases RNA polymerase II pausing mediated by a MED26 polypeptide.

**[00110]** Also provided is an agent or composition comprising an agent that increases RNA polymerase II pausing mediated by a MED26 polypeptide for use in a method of promoting erythroid differentiation in an erythroid precursor cell.

**[00111]** As used herein, a "carrier" includes pharmaceutically acceptable carriers, excipients, or stabilizers that are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Non-limiting examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

**[00112]** As used herein, the term "effective amount" or "therapeutically effective amount" of a substance is at least the minimum concentration required to effect a measurable improvement or prevention of a particular disorder. An effective amount herein may vary according to factors such as the disease state, age, sex, and weight of the patient, and the ability of the substance to elicit a desired response in the individual. An effective amount is also one in which any toxic or detrimental effects of the treatment are outweighed by the therapeutically beneficial effects. In reference to cancer, an effective amount comprises an amount sufficient to cause a tumor to shrink and/or to decrease the growth rate of the tumor (such as to suppress tumor growth) or to prevent or delay other unwanted cell proliferation in cancer. In some embodiments, an effective amount is an amount sufficient to delay development of cancer. In some embodiments, an effective amount is an amount sufficient to prevent or delay recurrence. In some embodiments, an effective amount is an amount sufficient to reduce recurrence rate in the individual. An effective amount can be administered in one or more administrations. An effective amount can be administered in one or more administrations. For purposes of this disclosure, an effective

amount of drug, compound, or pharmaceutical composition is an amount sufficient to accomplish prophylactic or therapeutic treatment either directly or indirectly. As is understood in the clinical context, an effective amount of a drug, compound, or pharmaceutical composition may or may not be achieved in conjunction with another drug, compound, or pharmaceutical composition. Thus, an "effective amount" may be considered in the context of administering one or more therapeutic agents, and a single agent may be considered to be given in an effective amount if, in conjunction with one or more other agents, a desirable result may be or is achieved.

## EXAMPLES

**[00113]** The following examples of the invention are to further illustrate the nature of the invention. It should be understood that the following examples do not limit the invention and that the scope of the invention is to be determined by the appended claims.

**Example 1. MED26 expression during erythropoiesis**

**[00114]** As the Mediator complex is crucial in the regulation of various transcription steps, its expression in a human CD34+ erythroid culture system was detected. CD34+ cells were purified from human cord blood (Cord Blood Bank of Beijing) using the MACS MicroBead kit (Miltenyi Biotec, Bergisch Gladbach, Germany). The two-phase erythroid differentiation protocol was modified from the previous study(38). The base medium contained IMDM (Gibco, Waltham, MA, USA), 5% human AB serum (Wokavi Biotech, Beijing, China), 10% FBS (Gibco), 10 ng/mL heparin (Sigma, St. Louis, MO, USA), 10 μg/mL insulin (Sigma), 2 mM L-Glutamine (Gibco), 3 IU/mL erythropoietin (Amgen, Thousand Oaks, CA, USA), and 300 μg/mL holo-Transferrin (Sigma). Phase-I medium was supplemented with 50 ng/mL Human Recombinant SCF (Stem Cell Technologies, Vancouver, BC, Canada) and 10 ng/mL Human Recombinant IL-3 (Stem Cell Technologies). Phase-II medium was supplemented with 50 ng/mL Human Recombinant SCF only. Cells were cultured in Phase-I medium for 8 days and then transferred to Phase-II medium for 6-8 days. All mouse and human cells were cultured in a 90% (v/v) humidified atmosphere with 5% (v/v) CO2 at 37 °C.

**[00115]** The protein level of most Mediator subunits diminished substantially in the terminal stages of erythroid differentiation, whereas that of MED26 remained detectable throughout unexpectedly (Fig.1A). Consistently, immunofluorescence imaging of primary erythroblasts isolated from E14.5 mouse fetal livers revealed that the MED1 level declines while MED26

remains detectable from early to late erythropoiesis (Data not shown). Given the relative abundance of MED26 in terminal erythropoiesis and the critical function of the Mediator complex in transcription regulation, MED26 may serve a unique function in late erythropoiesis.

**Example 2. MED26 exhibits biomolecular condensate formation**

**[00116]** During erythropoiesis, it was not only found that the level of most Mediator subunits including MED1 declined, but that of representative transcription condensate forming proteins, such as BRD4 and RPB1, also diminished substantially (Fig.1A, B). Taken together, these findings suggest a distinct composition of transcription condensates during terminal erythroid differentiation.

**[00117]** It was next tested whether MED26 has condensate forming capability, along with other Mediator subunits, by conducting the OptoDroplet assay. The OptoDroplet assay was conducted as previously described (35). Plasmids containing the indicated Mediator subunits were transfected with Lipo2000 reagent into 293T cells. Transfected 293T cells were cultured at 37 °C for 48 h. For blue light activation and imaging, the cells were imaged using two laser wavelengths, 488 nm and 568 nm, every 2 seconds. The proteins with phase separation capacity will aggregate when triggered by blue light (Fig.1C). This assay revealed that MED1, MED4, MED26, and MED28 had droplet-forming capability (Fig.1C) and structural evidence showing that large parts of these proteins are unstructured.

**[00118]** Whether MED26 exhibits droplet properties *in vivo* was examined using a fluorescence recovery after photobleaching (FRAP) assay. Briefly, the FRAP experiments were performed on a spinning disk microscope with a 63× oil objective. MED26 droplets were bleached for 30 cycles using a laser intensity of 80% at 480 nm (for GFP). The fluorescence recovery was recorded accordingly, and the fluorescence intensity of the photobleached area was normalized to the intensity of the unbleached area. The FRAP assay showed that enhanced green fluorescence protein fused with MED26 (EGFP-MED26) can form condensates in K562 erythroleukemia cells, and that the signal recovers promptly after photo-bleaching (Fig.1D).

**[00119]** To validate the intrinsic condensate forming capacity of MED26, the putative EGFP-MED26-IDR was expressed and purified. The DNA sequence of IDRs for the gene of interest was cloned into the pET28a prokaryotic expression vector, the backbone of which contains sequences encoding a 5' 6xHis-tag followed by either mCherry or EGFP and an artificially synthesized amino acid linker GAPGSAGSAAGGSG (SEQ ID NO: 18). The DNA sequence

was inserted in frame into the backbone and ended with a STOP codon. For protein expression, the plasmids were first transformed into Rosetta competent cells (Tsingke, Beijing, China); a fresh bacterial colony was inoculated into 15 mL LB media containing kanamycin and chloramphenicol and grown at 37 °C overnight. The bacteria were subsequently transferred to a 500 mL culture and incubated until the optical density (O.D.) reached approximately 0.3. The cells were pre-chilled to 16 °C before 1 mM IPTG induction and then grown overnight in a 16 °C centrifuge at 130 rpm. The bacterial pellets were collected and stored at -80 °C before further processing.

**[00120]** The bacterial pellets were resuspended in Buffer A (50 mM Tris pH 7.5, 500 mM NaCl) containing 10 mM imidazole, lysozyme, and a protease inhibitor cocktail. The cell resuspension was fully lysed on ice for approximately 30 min and sonicated (ten cycles of 15 seconds on, 60 seconds off) until lysates turned slightly clear. The lysates were centrifuged at 12,000g for 30 min to remove insoluble impurities. The supernatant was incubated with 1 mL pre-equilibrated Ni-NTA agarose at 4 °C for at least 1.5 h. The slurry was transferred into a column and washed with 15 volumes of Buffer A containing 10 mM imidazole. The purified protein was sequentially eluted with two volumes of Buffer A containing 50 mM imidazole, two volumes of Buffer A containing 100 mM imidazole, and three volumes of Buffer A containing 250 mM imidazole. The composition and purity of each fraction were analyzed using SDS-PAGE electrophoresis followed by Coomassie blue staining. The fraction containing the target protein was concentrated to a suitable volume and stored at -80 °C.

**[00121]** For the droplet formation assay with or without crowding agents, buffers containing 10% PEG8000 in phosphate-buffered saline (PBS), or low salt buffer containing 50 mM Tris-HCl (pH 7.5) and 25 mM NaCl without any crowding agents were used, respectively. Recombinant proteins were exchanged into the droplet formation buffer using Amicon Ultra centrifugal filters (30kDa MWCO, Millipore, Burlington, MA, USA). The protein solution was immediately loaded onto a homemade chamber comprising a glass slide with a coverslip attached using two parallel strips of double-sided tape. Slides were then imaged using an Andor confocal microscope with a 63× objective (Oxford Instruments, Abingdon, UK).

**[00122]** It was observed that EGFP-MED26-IDR can form phase-separated droplets in vitro, with or without crowding reagents (10% PEG8000) (FIG.1E and 1F). Using time-lapse imaging, fusion events of the MED26 droplets were captured, indicating their fluidic property (Data not

shown). The droplet formation assay with varying concentrations of EGFP-MED1-IDR, EGFP-MED26-IDR, and EGFP was next performed, which showed that MED26-IDR has a lower saturation concentration of phase separation than MED1-IDR (FIG.1E).

**[00123]** Moreover, MED26-IDR droplets are sensitive to 1,6-HD and high salt treatment, indicating that hydrophobic and electrostatic interactions contribute to the droplet formation (FIG. 1F). Collectively, the results demonstrated that MED26, a subunit in the core Mediator complex, can undergo phase separation in vitro and in vivo, and suggested that the transcription condensates switch to a "MED26 enriched form" during terminal erythroid differentiation.

**Example 3. MED26 plays a crucial role in hematopoietic development in mice and human**

**[00124]** To investigate the role of Med26 *in vivo*, *Med26* conditional knock-out (cKO) mice were constructed to examine its function in the hematopoietic system. To construct Med26 conditional knockout mice, CRISPR/Cas9 technology was used to modify the Med26 gene. Flox sequences were inserted into both sides of the Med26 gene loci. The brief process is as follows: sgRNA, Cas9, and flox sequence donors were microinjected into the fertilized eggs of C57BL/6JGpt mice. Fertilized eggs were transplanted to obtain positive F0 mice which were confirmed by PCR and sequencing. A stable F1 generation mouse model was obtained by mating positive F0 generation mice with C57BL/6JGpt mice. Mx1-iCre mice were also generated by CRISPR/Cas9 technology, the Mx1-iCre-polyA gene fragment was inserted into the H11 site of mice. The brief process is similar to Med26 conditional knockout mice. Med26 flox/flox; wt/wt (control) or Med26 flox/flox; Mx1-iCre/wt (cko) mice were generated by natural propagation. To induce iCre expression in vivo, 10μg g−1 polyinosinic-polycytidylic acid (pIpC-HMW, InvivoGen) was injected into mice. After one week, mice were sacrificed and the organs of interest were isolated for the following analysis.

**[00125]** Med26 genomic loci was recombined and its protein expression was down-regulated after Cre induction by pIpC injection (Fig.2A). The cKO mice showed smaller body size, whiter hair and worse survival rate (FIG.2B). Prussian blue staining of the spleen section showed more iron deposition in the cKO mice, indicating increased abnormal erythroblasts were removed by the spleen (Fig.2C). Complete blood count (CBC) analysis of peripheral blood showed that cKO mice had lower mean corpuscular volume (MCV) and mean corpuscular hemoglobin (MCH), but similar RBC, HGB, and hematocrit (Fig.2D), compared to the wildtype mice. Giemsa staining of the peripheral blood smear also showed the presence of irregular-shaped RBCs and scarce

reticulocytes (Fig.2E). Erythroid precursors (especially at the S3 stage) were absent in the bone marrow of Med26 cKO mice (Fig.2F), which phenocopies the cKO mice of *Gata1*, the master erythroid transcription factor (37).

**[00126]** The erythroid regeneration capacity of the cKO mice was next tested under phenylhydrazine (PHZ)-induced acute hemolytic anemia condition. Briefly, Med26 $^{flox/flox}$; wt/wt (control) or Med26 $^{flox/flox}$; Mx1-iCre/wt (cko) mice, 6–8 weeks of age, randomized by weight, were pretreated with pIpC (5μg g−1) for 1 day (day −1). On day 0, the mice were injected with phenylhydrazine (PHZ, Sigma) (60 mg kg−1). Whole blood samples were collected on each day of days 1–7 for complete blood count (CBC) analyses. Results showed that cKO mice were unable to recover after PHZ treatment (FIG.2G), consistent with much smaller spleen, scarce reticulocytes in the peripheral blood and fewer erythroblasts at the S3 stage in the bone marrow These results demonstrated that MED26 is essential for erythropoiesis in normal and stress conditions. The hematopoietic stem and progenitor cell (HSPC) and white blood cell (WBC) composition of cKO mice was also analyzed, which showed that LT-HSC, ST-HSC, MPP, CMP, GMP, MEP, B cell and neutrophils are impaired upon MED26 deficiency (Fig.2H and 2I).

**[00127]** In order to test the effects of knocking down MED26 in human cells, CD34+ cells were purified from human cord blood as described above. For gene knockdown in CD34+ cells, a miR-30-based shRNA vector was used; it was created by inserting the target gene into the *Eco*RI and *Xho*I sites as previously described(*60*). It was observed that knocking down MED26 substantially abrogated colony-forming capacity of human CD34+ HSPCs (FIGs. 3A-D). Collectively, it was demonstrated that Med26 is essential for mice and human erythropoiesis.

**Example 4. Condensate forming capacity of MED26 is essential for erythropoiesis**

**[00128]** In order to address the function of MED26 during erythroid development, a gain-of-function assay in the primary human erythroid culture system was conducted. For gene overexpression in CD34+ cells, the coding sequence of the gene was amplified using PCR and cloned into the *Bam*HI and *Not*I sites of the MSCV-3XFlag-T2A-copGFP vector. Results showed that MED26 over-expression promoted erythroid differentiation, but inhibited enucleation (Fig.4A-C).

**[00129]** Moreover, to examine whether the function of MED26 in erythropoiesis is mediated by its phase-separation capacity, the OptoDroplet assay was first used to assess the function of MED26 truncations (Fig.4D). The results showed that the truncations containing amino acids

(a.a.) 88-480, the putative IDR region, have the aggregate forming capacity (Fig.4D). Interestingly, a.a. 1-480 formed more conspicuous aggregates than all other truncations, which suggests that a.a. 1-87 can promote the phase separation capacity of the MED26 IDR. In vitro assays further validated that a.a. 88-480 of MED26, but not a.a. 1-87 or 480-600, had phase separation capacity (Fig.4E). Previous reports showed that a.a. 1-87 is the TFIIS domain and that a.a. 480-600 contains the surface that interacts with the remaining Mediator complex(36, 39). Then these three segments of MED26 were overexpressed in human CD34+ HSPCs, which showed that a.a. 88-480 alone can sufficiently promote erythroid differentiation relative to the full-length MED26 (Fig.4F). Collectively, these results indicate that the phase-separation capacity of MED26 is a key regulating factor to promote erythroid differentiation.

**Example 5. MED26-enriched transcription condensates are preferentially associated with RNA polymerase II pausing**

**[00130]** To further dissect the molecular basis underlying MED26 function, the CUT&Tag signals of MED26 were analyzed, and compared to those of MED1, which is often regarded as a representative subunit of Mediator, in human CD34+ erythroid culture cells on Day 4. The CUT&Tag assay was performed as previously described (59) and followed the manufacturer's instructions for the kit (Vazyme, Nanjing, China). Briefly, 100,000 cells were freshly collected and captured using Concanavalin A (ConA) beads at room temperature. The cell/ConA beads complex was permeabilized and incubated with the primary antibody (at 1:100 dilution) overnight at 4 °C. The secondary antibody was added to the solution and incubated at room temperature for 1 h. The secondary antibody was washed off prior to adding pA/G-Tn5 to the cell suspension and then incubated at room temperature for 1 h. The adapters were then inserted into the Tn5-tagged genome using the TruPrep Tagment Buffer L (TTBL). The genomic DNA was extracted with DNA beads, and the tagged DNA fragments were amplified using the Next-Generation Sequencing (NGS) adapters with 16 cycles of PCR. The adaptor dimer was removed from the PCR products with DNA Clean Beads (Vazyme, Nanjing, China). The library was quantified before sequencing.

**[00131]** Unexpectedly, results showed that ~60% of chromatin sites with MED1 or MED26 occupancy do not colocalize (Fig.5A). Interestingly, MED26 colocalizes with GATA1 and GATA2 better than MED1, which suggests potential functional association between MED26 and GATA factors. The presence of MED1- or MED26-enriched condensates were observed in the

primary human erythroblasts, as well as other cell types. Then the signal ratio of MED26 to MED1 was calculated at transcription start sites (TSS) of their occupancy loci. While a large fraction of MED1 and/or MED26-occupied genes have a relatively constant MED26/MED1 signal ratio, a considerable number of genes have differential MED26/MED1 ratios, indicating the chromatin occupancy of MED26 and MED1 do not always intercorrelate. The maximum 10% ratio were defined as the "MED26-enriched chromatin sites" and the minimum 10% ratio as the "MED26-poor chromatin sites" (Fig.5B-C).

**[00132]** To elucidate how the enrichment of MED26 impacts transcription, a CUT&Tag assay of Pol II was conducted and the precision nuclear run-on sequencing (PRO-seq) in human CD34+ erythroid culture cells on Day 4. Results showed Pol II has more signals spanning through the gene body at the MED26-poor chromatin sites, whereas Pol II often has a distinct enrichment at the TSS region, transcribing a higher percentage of short RNAs at the MED26-enriched chromatin sites, indicating Pol II pausing (Fig.5D-E). To better characterize this observation, the Pausing Index (PI) was calculated, which is defined by the ratio of "the read counts at the TSS region" to "the read counts at the gene body region" after normalizing both read counts by the length of the genomic region (40). Analysis revealed that the MED26-enriched loci have a significantly higher PI ($p < 2.2 \times 10^{-16}$), longer gene length ($p = 4.092 \times 10^{-10}$), and similar exon numbers ($p = 0.09774$) compared to the MED26-poor loci (Fig.5F). The same conclusion can be drawn from CUT&Tag assays of Ser5 phosphorylation of the RNA polymerase II carboxy-terminal domain (Rpb1-S5) and Rpb1-S2 (Fig.5F). Based on the overall quantification, it was found that Pol II in MED26-enriched transcription condensates transcribes genes with a higher pausing index, lower elongation efficiency and a higher percentage of short RNAs.

**Example 6. MED26 recruits pausing factors to form biomolecular condensates**

**[00133]** To examine how MED26 might mediate transcription pausing, immunoprecipitation coupled with Mass Spectrometry (IP-MS) was performed to identify interacting proteins of MED1 and MED26. HEK293T cells transfected with indicated bait constructs (3XFlag, 3XFlag-MED1, or 3XFlag-MED26) were collected, resuspended in lysis buffer (Beyotime, China), and incubated on ice for 30 min. Soluble protein complexes were collected by centrifuging at 15,000*g* for 15 min; the supernatants were then incubated with anti-Flag agarose beads (Sigma) for 1.5 h at 4 °C. After five washes with lysis buffer, the protein complexes were eluted with 100 μg/mL 3XFlag peptide (Beyotime, China), denatured with 2× SDS sample buffer, and analyzed

using SDS-PAGE electrophoresis followed by Coomassie blue staining. The PAGE gel band of the expected molecular weight for the target protein was cut into suitable size, digested with trypsin, and analyzed on a Q Exactive Plus Mass Spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) to identify the interacting proteins. The MS data were aligned with the Human Reviewed Swiss-Port database using the Proteome Discoverer 2.2 software.
The NSAFs were calculated for each detected protein(*32, 61-63*). The NSAF for a protein k is proportional to the amount of the protein present in the sample and is calculated using the following formula:

$$(NSAF)_k = \frac{(S_pC/L)_k}{\sum_{i=1}^{N}(S_pC/L)_i}$$

where SpC = spectral count, L = protein length in amino acids, and i = all proteins detected in the MudPIT runs.

**[00134]** While MED26 and MED1 similarly co-immunoprecipitated with the majority of other Mediator subunits (the head, middle, and tail modules) and the elongation complex, MED26 interacted much less with the CDK8 kinase module than MED1, consistent with a previous study(30). Notably, MED26 interacted with many more transcription pausing factors (NELF, DSIF, and PAF complexes) than MED1, which supports the notion that MED26 associates with pausing (FIGs.6A-B). To determine whether the recruitment of pausing factors depends on the phase separation of MED26, mCherry fused with the disordered region of PAF1 was purified (FIG.6C). The in vitro droplet formation assay not only shows that a.a.1-480 of MED26 can recruit pausing factor PAF1 (Fig. 6D), but also suggests that the presence of pausing factors reciprocally promotes condensate formation of MED26. Previous studies show that a.a. 1-87 of MED26 contain a TFIIS domain, which can interact with TFIIS interacting motifs (TIM) that ubiquitously exist in disordered regions of transcription regulators(41). Indeed, the PAF1 complex subunits PAF1, LEO1, and CTR9 all contain TIM sequences. Therefore, the TFIIS domain and the IDR domain of MED26 may both be key segments to promote condensate formation. To investigate whether the IDR of MED26 contributes to the recruitment of pausing factors, EGFP fused with various truncations of MED26 were constructed and the recruitment of PAF1-mCherry was tested. The results demonstrate that MED26's ability to recruit PAF1 is contingent on its IDR, which cannot be substituted by the FUS IDR domain (FIG. 6E). These

observations imply that MED26's IDR domain serves a distinctive role in recruiting PAF1. Taken together, both the IDR and TFIIS domains of MED26 are necessary for the recruitment of pausing factors, and the presence of pausing factors enhances the condensate formation of MED26.

**[00135]** To address how the enrichment of MED26 impacts transcription during erythropoiesis, CUT&Tag assays were used to analyze the ratio of MED26 to MED1 chromatin occupancy on Day 4 and Day 16 of primary human CD34+ derived erythroblasts. Results revealed that the ratio of MED26 to MED1 increases at both erythroid and non-erythroid genes during differentiation, and the ratio at erythroid genes on Day 16 has a more marked increase relative to non-erythroid genes (FIG.6F-G).

**[00136]** Then the PRO-seq was performed on Day 4 or Day 16 ex vivo cultured erythroblasts. PRO-seq libraries were prepared as described previously (*58*). Briefly, $2\times10^7$ Day4 or Day16 ex vivo cultured permeabilized CD34 cell nuclei were added to 2 × Nuclear Run-On (NRO) reaction mixture and incubated for 3 min at 37°C. Nascent RNA was extracted and fragmented by base hydrolysis in 0.2 N NaOH on ice for 10 min, and neutralized by adding 1 × volume of 1 M Tris-HCl pH 6.8. Fragmented nascent RNA was exchanged buffer and purified using streptavidin beads, ligated with reverse 3' RNA adapter VRA3 (5'p-GAUCGUCGGACUGUAGAACUCUGAAC (SEQ ID NO: 19)-/3'inverted dT/), and the products were enriched by a second round of streptavidin enrichment. Then for 5' end repair, the RNA products were treated with RppH (Thermo Fisher) and polynucleotide kinase (PNK, NEB). 5' repaired RNA was ligated to reverse 5' RNA adaptor VRA5 (5'-CUGAACAAGCAGAAGACGGCAUACGA (SEQ ID NO: 20)-3') before the third round of streptavidin bead enrichment. RNA was reverse transcribed using RT primer (5'AATGATACGGCGACCACCGACAGGTTCAGAGTTCTACAGTCCGA (SEQ ID NO: 21)-3'). The cDNA product was amplified for 18 cycles and the library larger than 150 bp (insert > 70 bp) is purified from PAGE gel. The library was quantified before NGS.

**[00137]** Results revealed that the PI increases at both non-erythroid genes ($p < 2.2 \times 10^{-16}$) and erythroid genes ($p < 2.2 \times 10^{-16}$), along with the increase of the MED26 to MED1 ratio during erythroid development (FIGs.6G-H). To test whether the increase of transcription pausing promotes erythropoiesis, 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) was added, which represses transcription elongation by inhibiting CDK9(42), to Day 6 ex vivo cultured

human erythroblasts. The addition of DRB increased the number of $CD71^+$ $CD235a^+$ erythroblasts compared to the mock group (Fig.6I). Collectively, it was found that MED26 enrichment is associated with increased transcription pausing during terminal erythroid stages, which can promote erythroid development.

**Example 7. mRNA Delivery Mediated by MED26 coacervates**

**[00138]** mRNA therapy holds promising potential as a revolutionary approach in the field of medicine. Its ability to leverage the body's natural cellular machinery for protein production offers targeted and precise treatment for various diseases. Therefore, the potential of MED26 condensates to function as intracellular transport for mRNA was evaluated.

**[00139]** Condensates of MED26 (1-480 amino acids; SEQ ID NO: 2) and mRNA were prepared by adding them in $ddH_2O$ and gently mixing for 30-60 seconds. The phase separation diagram was determined through microscopic observation for a range of MED26 (1-480 amino acids; SEQ ID NO: 2) concentrations (0.1-100 μM) and mRNA concentrations (0.01-0.6 μM).

**[00140]** PEG was dissolved in $ddH_2O$ with a series of concentrations (0, 2, 5,10, 15, 20%, wt/v). MED26 (1-480 amino acids; SEQ ID NO: 2) (10 μM) and EGFP mRNA (0.1 μM) were introduced into the polyethylene glycol (PEG) solution and gently mixed for 30-60 seconds to form coacervates. $1\times10^5$ Jurkat cells were suspended in 20 μL Opti-MEM and combined with 40 μL coacervate solution. After a 2-hour incubation in a 96-well plate at 37℃ with 5% $CO_2$, 200 μL cell medium (RPMI-1640, 10% FBS, antibiotics) was added and the cells were cultured for 18 hours.

**[00141]** For adherent cell, $5\times10^4$ 293T cells were seeded in a 96-well plate one day before treatment. Following the removal of the culture medium and a single wash with Opti-MEM to clean up FBS, 20 μL Opti-MEM and 40 μL coacervate (formula of 10% PEG) solution were added for a 2-hour incubation. Subsequently, 200 μL cell medium (DMEM, 10% FBS, antibiotics) was supplemented, and the cells were cultured for 18 hours before testing. The expression efficiency was assessed by flow cytometry (Cytoflex LX). GFP fluorescence images were captured using (TSKON TS2-FL).

**[00142]** Treating MED26-mRNA condensates with the mouse monocyte/ macrophage-like cell line RAW264.7 revealed various stages of cell uptake captured through electron micrography. This process displayed characteristics of micropinocytosis and phagocytosis, involving stages

such as no-touch, attachment to the cell surface, coacervate wrapping by flat membrane protrusion, and partial membrane engulfment (FIG. 7).

**[00143]** PEG is widely employed to enhance LLPS as a crowding agent and dehydration inducer. It was observed that coacervates formed in the PEG solution demonstrated higher efficiency in cell entry in a concentration-dependent manner by detecting Cy5-labeled mRNA (FIG. 8). The relationship between PEG concentration during the MED26-mRNA coacervate generation and expression efficiency was also investigated. FIG. 8 that higher expression level of the protein encoded by the mRNA was achieved when the coacervate was assembled in PEG solutions, and the high expression was achieved in 10% or higher PEG solution. MED26-EGFP mRNA coacervates efficiently transfected 293T, expressing EGFP mRNA with comparable efficiency to Lipofectamine MessengerMAX reagent (FIGs. 9A-B). Significantly, MED26 coacervates exhibited a notable capacity for delivering mRNA into Jurkat cells, surpassing the performance of Lipofectamine MessengerMAX reagent (FIGs. 9A-B).

**[00144]** The primary function of the MED26 protein is to participate in gene transcription within the cell nucleus. To retain its delivery capability while losing its function, a series of truncated forms of the MED26 protein were designed and purified (FIG. 10A). All truncations tested exhibited the ability to form coacervates with mRNA. Therefore, their delivery efficiency in Jurkat cells was compared and screened (FIG. 10B). Considering the removal of the nuclear localization sequence and expression efficiency, the fragment of amino acids 135-480 of MED26, i.e., MED26 (135-480; SEQ ID NO: 12) protein, for the subsequent experiments was chosen. To determine the optimal concentration of MED26 (135-480; SEQ ID NO: 12) protein, the delivery efficiency of various protein concentrations were assessed with 0.1 μM of mRNA, and 10 μM of protein was found to be sufficient to achieve high mRNA delivery efficiency in Jurkat cells (FIG. 10C). Coacervates produced by higher MED26 concentration also successfully delivered mRNA into Jurkat cells (Fig.10D).

**[00145]** Moreover, MED26 coacervates were not confined to carrying a single mRNA; confocal observation revealed complete co-localization of two mRNAs within MED26 coacervates (data not shown). FACS analysis and confocal observations confirmed a high correlation in expression between EGFP mRNA and tdTomato mRNA in Jurkat cells (FIG. 11) and in 293T cells.

**Example 8. Internalization mechanism and duration of MED26 coacervates**

**[00146]** Confocal observations indicated that MC38 cells exhibited deformation and extended pseudopodia to uptake MED26 coacervates, a characteristic feature of macropinocytosis (FIG. 12A). To gain deeper insights into the internalization mechanism of MED26 coacervates, their localization in relation to early endosomes was investigated. $5\times10^4$ 293T cells were seeded in the confocal dish (solarbio) and transfected for 16 hours with plasmids expressing early endosome-GFP (Rab5a, BacMam 2.0, CellLight, Thermo Fisher Scientific). Subsequently, cells were treated with MED26 (135-480; SEQ ID NO; 12) -luciferase mRNA (Cy5 labeled) coacervates for 30 minutes, followed by a medium change to DMEM (10% FBS, antibiotics). The images were captured using laser scanning confocal microscopy (Zeiss LSM980).

**[00147]** After 20 minutes post-transfection with MED26 coacervates, the majority of coacervate signals in 293T cells did not co-localize with Rab5 signals (FIG. 12A). Even after an additional 40 minutes, no significant co-localization was observed (FIG. 12B), suggesting that the internalization mechanism does not involve endosome.

**[00148]** To further confirm that MED26 coacervates do not get trapped in endosomal compartments, 293T cells were stained with LysoTracker, revealing acidic organelles like lysosomes, 3 hours post-transfection. The images in FIG. 12C taken at 4 hours post-transfection, clearly demonstrate that almost all coacervates did not co-localize with lysosomes and maintained a high fluorescent intensity. These findings strongly indicate that the internalization of MED26 coacervates, characterized by micro-scale and liquid-like droplets, is distinctly different from endocytosis.

**[00149]** Since there is no entrapment in the endosome and no necessity to escape from it, the GFP signal became detectable as early as 30 minutes post-transfection, reaching its peak before 8 hours with a quite high expression ratio in Jurkat cells (FIG. 12C). The intensity of GFP fluorescence reached its maximum at 1day post-transfection, followed by a decrease. Concurrently, mRNA level in cells rapidly declined within 1 day. MED26 protein was scarcely detectable at 1 day post-transfection (Figure 12D). The data showed that the MED26 protein is not sustained for an extended period within the cells.

**Example 9. siRNA Delivery and Gene Knockdown Mediated by MED26 coacervates**

**[00150]** In this study, the effective delivery of siRNA through MED26 coacervates was studied. For the single-gene siRNA knockdown experiment, 293T-GFP cells expressing stable GFP were seeded at a density of $2.5\times10^5$ cells per well in a 24-well cell culture plate one day prior. A

coacervate mixture containing 10 μM MED26 (135-480; SEQ ID NO: 12), 10% PEG, and 1 μM siRNA was prepared by thorough mixing for 30-60 seconds. The existing culture medium in each well was removed, and 200 μL of the coacervate mixture, combined with 100 μL of opti-MEM medium, was added to each well. The plate was then incubated stationary in a cell culture incubator for 2 hours. Post-transfection, 1 mL of complete DMEM growth medium was added to each well, and the plate was further incubated at 37℃ in the incubator for 72 hours to assess the siRNA's knockdown effect on the target gene.

**[00151]** For multi-gene co-knockdown experiments, siRNAs targeting different genes were thoroughly mixed, and a siRNA mixture with a final concentration of 1 μM was used for coacervate preparation. The remaining transfection conditions remained unchanged.

**[00152]** At 72 hours post-transfection with siGFP-loaded MED26 coacervates, GFP fluorescent intensity and GFP mRNA levels markedly decreased in GFP-stable expression 293T cells (data not shown). Additionally, other siRNA targets were validated in 293T-GFP cells, all demonstrating significant knockdown efficiency (FIG. 13A). Due to MED26's high loading capacity, the functionality of multiple gene knockdown simultaneously was assessed. Transfecting cells with quadruple siRNAs-loaded MED26 coacervates exhibited substantial targeted gene knockdown, indicating the potential for concurrent multi-gene knockdown using MED26 coacervates (FIG. 13B).

**Example 10. DNA Delivery and pDNA based genomic editing Mediated by MED26 coacervates**

**[00153]** CRISPR/Cas9-mediated gene knockout using plasmid DNA (pDNA) is both accessible and cost-effective. Initially, pDNA delivery was assessed through MED26 coacervates. MED26-GFP pDNA coacervates successfully transfected Jurkat cells, surpassing the capabilities of Lipofectamine 3000 (FIG. 14A). The transfection efficiency was comparable in 293T cells when delivering a small-sized plasmid, but notably, MED26 coacervates demonstrated enhanced performance in delivering large-sized plasmids (FIG. 14B). Furthermore, MED26 coacervates outperformed Lipofectamine 3000 when co-transfecting cells with two pDNAs simultaneously (FIGS. 14C-D). For evaluating transfection efficiency, an all-in-one plasmid encoding the sgRNA and Cas9-tdTomato fusion protein sequence was selected for MED26 coacervates (FIG. 14E). Subsequently, the MED26 protein and pDNA formulation was optimized, determining that 10 μM MED26 and 20 μg/μL pDNA were the optimal concentrations.

The gene editing efficiency facilitated by MED26 coacervates was next assessed. Jurkat cells were exposed to CRISPR/Cas9 pDNA-loaded MED26 coacervates targeting DNA methyltransferase 1 (DNMT1), hemoglobin subunit beta (HBB), hypoxanthine phosphoribosyl transferase 1 (HPRT1), and a pDNA mixture of these three targets. To assess the efficiency of genomic editing, cells were treated with MED26 coacervate carrying the all-in-one plasmid (with gRNA for a specific target). After 48 hours of transfection, cells were harvested, and the genomic DNA was extracted by QuickExtract™ DNA Extraction Solution (Lucigen). The target genomic locus was amplified by nested PCR using PrimeSTAR® Max DNA Polymerase (Takara) and primers listed in Table 1 and purification was performed using the TaKaRa MiniBEST Agarose Gel DNA Extraction Kit (Takara). Subsequently, 200 ng of PCR products underwent the T7EN1 assay to evaluate the indel frequency following recommended protocols (Vazyme Biotech). The digested products were analyzed using 2% agarose gels and captured with a gel imaging system (Tanon). Indel percentages were measured by ImageJ software and calculated with the following formula:

Indel (%) = 100 × [1- (1- fraction cleaved)$^{1/2}$], where the fraction cleaved is defined as the sum of each digested product intensity divided by the sum of each digested product intensity and undigested product intensity.

**[00154]** At 48 hours post-transfection, approximately 73.9%, 67.9%, 59.6%, and 58.1% of cells expressed tdTomato, indicating the proportions of cells expressing the target sgRNAs and Cas9 (FIG. 14F).

**[00155]** The analyzed indel efficiencies were 23.8%, 17.55%, and 34.79% at the DNMT1, HBB, and HPRT1 loci in the single-targeted group, respectively (FIG. 14G). Notably, genome editing resulted in 16.43%, 10.29%, and 28.05% at the DNMT1, HBB, and HPRT1 loci in the triple-targeted sample, suggesting the significant potential for simultaneous multiple-gene editing (data not shown). Due to its capacity to carry larger and more plasmids, MED26 coacervates demonstrated higher efficiency of plasmid expression ratios compared to Lipofectamine 3000 reagent in 293T cells (FIG. 14H). Consequently, MED26 coacervates mediated higher genome editing efficiency at the DNMT1, HBB, HPRT1, and multiple-targeted loci (FIG. 14I).

**Table 1. Primer Sequences**

| sgRNA sequence constructed in Crispr/Cas9 pDNA |
| --- |
| |

| Gene | sgRNA（oligo1） | sgRNA（oligo2） |
|---|---|---|
| DNMT1 | CACCGTAGGAGATCTCCAGTGCCG (SEQ ID NO:22) | AAACCGGCACTGGAGATCTCCTAC (SEQ ID NO:23) |
| HBB | CACCGTAACGGCAGACTTCTCCTC (SEQ ID NO:24) | AAACGAGGAGAAGTCTGCCGTTAC (SEQ ID NO:25) |
| HPRT1 | CACCGCATTTCTCAGTCCTAAACA (SEQ ID NO:26) | AAACTGTTTAGGACTGAGAAATGC (SEQ ID NO:27) |
| **Primers to detect gene editing efficiency byT7 enzyme digestion** | | |
| Gene | Forward primer | Reverse primer |
| HBB | AACTCCTAAGCCAGTGCCAGAAGAG (SEQ ID NO:28) | CAGGCCATCACTAAAGGCACCGAGC (SEQ ID NO:29) |
| HPRT1 | CCTCATGGACTAATTATGGACAGGT (SEQ ID NO:30) | AGCTGAACAAGTACCAAACATGTAA (SEQ ID NO:31) |
| DNMT1 | CTGTGATCTCGGGAGAAGGC (SEQ ID NO:32) | GCAGTGGCTTGACACCATTC (SEQ ID NO: 33) |
| **siRNA sequences for coacervate delivery in 293T cells** | | |
| Gene | Sense（5'-3'） | Antisense（5'-3'） |
| HPRT1 | GATTGTTGTTTAACTTGTATT (SEQ ID NO:34) | TACAAGTTAAACAACAATCTT (SEQ ID NO:35) |
| GSK3B | GCATTTATCGTTAACCTAATT (SEQ ID NO:36) | TTAGGTTAACGATAAATGCTT (SEQ ID NO:37) |
| ITGB1 | GGATATTACTCAGATCCAATT (SEQ ID NO:38) | TTGGATCTGAGTAATATCCTT (SEQ ID NO:39) |
| IGF2R | CGCAGGTAACGATGGGAAATT (SEQ ID NO:40) | TTTCCCATCGTTACCTGCGTT (SEQ ID NO:41) |
| RB1 | CGTGTAAATTCTACTGCAATT (SEQ ID NO:42) | TTGCAGTAGAATTTACACGTT (SEQ ID NO:43) |
| FGFR1 | GAGATTTACCCATCGGGTATT (SEQ ID NO:44) | TACCCGATGGGTAAATCTCTT (SEQ ID NO:45) |
| BAL11A | GGCAGCCTCTGCTTAGAAATT (SEQ ID NO:46) | TTTCTAAGCAGAGGCTGCCTT (SEQ ID NO:47) |

| | | |
|---|---|---|
| GFP | ACAGCCACATGCACTTCAATT (SEQ ID NO:48) | TTGAAGTGCATGTGGCTGTTT (SEQ ID NO:49) |
| SiNC | TTCTCCGAACGTGTCACGTTT (SEQ ID NO:50) | ACGTGACACGTTCGGAGAATT (SEQ ID NO:51) |
| **Primers to detect siRNA delivery knockdown efficiency** | | |
| Gene | Forward primer | Reverse primer |
| HPRT1 | CCTGGCGTCGTGATTAGTGAT (SEQ ID NO:52) | AGACGTTCAGTCCTGTCCATAA (SEQ ID NO:53) |
| GSK3B | AGACGCTCCCTGTGATTTATGT (SEQ ID NO:54) | CCGATGGCAGATTCCAAAGG (SEQ ID NO:55) |
| ITGB1 | CCTACTTCTGCACGATGTGATG (SEQ ID NO:56) | CCTTTGCTACGGTTGGTTACATT (SEQ ID NO:57) |
| IGF2R | CTGCCGCTATGAAATTGAGTGG (SEQ ID NO:58) | CGCCGCTCAGAGAACAAGTT (SEQ ID NO:59) |
| RB1 | CTCTCGTCAGGCTTGAGTTTG (SEQ ID NO:60) | GACATCTCATCTAGGTCAACTGC (SEQ ID NO:61) |
| FGFR1 | CCCGTAGCTCCATATTGGACA (SEQ ID NO:62) | TTTGCCATTTTTCAACCAGCG (SEQ ID NO:63) |

**[00156]** Our design of an all-in-one Crispr/Cas9 pDNA successfully transfected three targets into Jurkat cells with high transfection percentage. Significantly, the MED26 coacervate can allow multiple gene edits through a single-step transfection, particularly in hard-to-transfect Jurkat cells, without the need for Cas9 protein or gRNA. Moreover, these coacervates can combine different gene categories, such as pDNA and siRNA, in a single coacervate (data not shown), offering more options for multi-target drug discovery.

REFERENCES

1. P. Wong et al., Gene induction and repression during terminal erythropoiesis are mediated by distinct epigenetic changes. Blood 118, e128-138 (2011).

2. D. Hnisz, K. Shrinivas, R. A. Young, A. K. Chakraborty, P. A. Sharp, A Phase Separation Model for Transcriptional Control. Cell 169, 13-23 (2017).

3. P. Cramer, Organization and regulation of gene transcription. Nature 573, 45-54 (2019).

4. J. Palis, Primitive and definitive erythropoiesis in mammals. Front Physiol 5, 3 (2014).

5. S. M. Hattangadi, P. Wong, L. Zhang, J. Flygare, H. F. Lodish, From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. Blood 118, 6258-6268 (2011).

6. H. Lodish, J. Flygare, S. Chou, From stem cell to erythroblast: regulation of red cell production at multiple levels by multiple hormones. IUBMB Life 62, 492-496 (2010).

7. Y. Mei, Y. Liu, P. Ji, Understanding terminal erythropoiesis: An update on chromatin condensation, enucleation, and reticulocyte maturation. Blood Rev 46, 100740 (2021).

8. M. Moras, S. D. Lefevre, M. A. Ostuni, From Erythroblasts to Mature Red Blood Cells: Organelle Clearance in Mammals. Front Physiol 8, 1076 (2017).

9. P. Valent et al., Normal and pathological erythropoiesis in adults: from gene regulation to targeted treatment concepts. Haematologica 103, 1593-1603 (2018).

10. Z. C. Murphy et al., Regulation of RNA polymerase II activity is essential for terminal erythroid maturation. Blood 138, 1740-1756 (2021).

11. J. Malik, J. A. Lillis, T. Couch, M. Getman, L. A. Steiner, The Methyltransferase Setd8 Is Essential for Erythroblast Survival and Maturation. Cell Rep 21, 2376-2383 (2017).

12. P. Kapoor-Vazirani, J. D. Kagey, P. M. Vertino, SUV420H2-mediated H4K20 trimethylation enforces RNA polymerase II promoter-proximal pausing by blocking hMOF-dependent H4K16 acetylation. Mol Cell Biol 31, 1594-1609 (2011).

13. S. Bottardi et al., Ikaros interacts with P-TEFb and cooperates with GATA-1 to enhance transcription elongation. Nucleic Acids Res 39, 3505-3519 (2011).

14. R. Kurita et al., Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. PLoS One 8, e59890 (2013).

15. P. Li et al., Phase transitions in the assembly of multivalent signalling proteins. Nature 483, 336-340 (2012).

16. S. Boeynaems et al., Protein Phase Separation: A New Phase in Cell Biology. Trends Cell Biol 28, 420-435 (2018).
17. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol 18, 285-298 (2017).
18. M. Boehning et al., RNA polymerase II clustering through carboxy-terminal domain phase separation. Nat Struct Mol Biol 25, 833-840 (2018).
19. Y. E. Guo et al., Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. Nature 572, 543-548 (2019).
20. H. Lu et al., Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. Nature 558, 318-323 (2018).
21. B. R. Sabari et al., Coactivator condensation at super-enhancers links phase separation and gene control. Science 361, (2018).
22. A. Boija et al., Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell 175, 1842-1855 e1816 (2018).
23. J. Soutourina, Transcription regulation by the Mediator complex. Nat Rev Mol Cell Biol 19, 262-274 (2018).
24. S. Sato et al., A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. Mol Cell 14, 685-691 (2004).
25. L. El Khattabi et al., A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. Cell 178, 1145-1158 e1120 (2019).
26. J. W. Yin, G. Wang, The Mediator complex: a master coordinator of transcription and cell lineage development. Development 141, 977-987 (2014).
27. M. Stumpf et al., The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220. Proc Natl Acad Sci U S A 103, 18504-18509 (2006).
28. M. Stumpf et al., Specific erythroid-lineage defect in mice conditionally deficient for Mediator subunit Med1. Proc Natl Acad Sci U S A 107, 21541-21546 (2010).
29. S. I. Kim, E. H. Bresnick, Transcriptional control of erythropoiesis: emerging mechanisms and principles. Oncogene 26, 6777-6794 (2007).
30. S. Malik, R. G. Roeder, Dynamic regulation of pol II transcription by the mammalian Mediator complex. Trends Biochem Sci 30, 256-263 (2005).

31. H. Takahashi et al., The role of Mediator and Little Elongation Complex in transcription termination. Nat Commun 11, 1063 (2020).
32. H. Takahashi et al., Human mediator subunit MED26 functions as a docking site for transcription elongation factors. Cell 146, 92-104 (2011).
33. H. Takahashi et al., MED26 regulates the transcription of snRNA genes through the recruitment of little elongation complex. Nat Commun 6, 5941 (2015).
34. H. M. Bourbon, Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. Nucleic Acids Res 36, 3993-4008 (2008).
35. Y. Shin et al., optoDroplets—Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. Cell 168, 159-171 e114 (2017).
36. X. Chen et al., Structures of the human Mediator and Mediator-bound preinitiation complex. Science 372, (2021).
37. L. Gutierrez et al., Ablation of Gata1 in adult mice results in aplastic crisis, revealing its essential role in steady-state and stress erythropoiesis. Blood 111, 4375-4385 (2008).
38. H. Y. Lee et al., PPAR-alpha and glucocorticoid receptor synergize to promote erythroid progenitor self-renewal. Nature 522, 474-477 (2015).
39. R. Peruzzini et al., 1H, 15N and 13C assignments of the N-terminal domain of the Mediator complex subunit MED26. Biomol NMR Assign 10, 233-236 (2016).
40. D. S. Day et al., Comprehensive analysis of promoter-proximal RNA polymerase II pausing across mammalian cell types. Genome Biol 17, 120 (2016).
41. K. Cermakova et al., A ubiquitous disordered protein interaction module orchestrates transcription elongation. Science 374, 1113-1121 (2021).
42. K. Yankulov, K. Yamashita, R. Roy, J. M. Egly, D. L. Bentley, The transcriptional elongation inhibitor 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits transcription factor IIH-associated protein kinase. J Biol Chem 270, 23922-23925 (1995).
43. B. Gaertner, J. Zeitlinger, RNA polymerase II pausing during development. Development 141, 1179-1183 (2014).
44. S. Alberti, A. Gladfelter, T. Mittag, Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell 176, 419-434 (2019).
45. A. G. Larson et al., Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. Nature 547, 236-240 (2017).

46. A. R. Strom et al., Phase separation drives heterochromatin domain formation. Nature 547, 241-245 (2017).
47. M. Yu et al., Visualizing the disordered nuclear transport machinery in situ. Nature, (2023).
48. A. Molliex et al., Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell 163, 123-133 (2015).
49. D. S. W. Protter, R. Parker, Principles and Properties of Stress Granules. Trends Cell Biol 26, 668-679 (2016).
50. M. Boehning et al., RNA polymerase II clustering through carboxy-terminal domain phase separation. Nature Structural & Molecular Biology 25, 833-+ (2018).
51. X. Han et al., Roles of the BRD4 short isoform in phase separation and active gene transcription. Nat Struct Mol Biol 27, 333-341 (2020).
52. F. X. Chen, E. R. Smith, A. Shilatifard, Born to run: control of transcription elongation by RNA polymerase II. Nat Rev Mol Cell Biol 19, 464-478 (2018).
53. X. Bai et al., TIF1gamma controls erythroid cell fate by regulating transcription elongation. Cell 142, 133-143 (2010).
54. A. A. Perreault, J. D. Brown, B. J. Venters, Erythropoietin Regulates Transcription and YY1 Dynamics in a Pre-established Chromatin Architecture. iScience 23, 101583 (2020).
55. D. J. Taatjes, A. M. Naar, F. Andel, 3rd, E. Nogales, R. Tjian, Structure, function, and activator-induced conformations of the CRSP coactivator. Science 295, 1058-1062 (2002).
56. D. J. Taatjes, R. Tjian, Structure and function of CRSP/Med2; a promoter-selective transcriptional coactivator complex. Mol Cell 14, 675-683 (2004).
57. F. A. Ran et al., Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281-2308 (2013).
58. D. B. Mahat et al., Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). Nat Protoc 11, 1455-1476 (2016).
59. H. S. Kaya-Okur et al., CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 10, 1930 (2019).
60. C. Fellmann et al., An optimized microRNA backbone for effective single-copy RNAi. Cell Rep 5, 1704-1713 (2013).

61. L. Florens et al., Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. Methods 40, 303-311 (2006).
62. Y. Zhang, Z. Wen, M. P. Washburn, L. Florens, Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. Anal Chem 82, 2272-2281 (2010).
63. A. C. Paoletti et al., Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. Proc Natl Acad Sci U S A 103, 18928-18933 (2006).
64. S. Pott, J. D. Lieb, What are super-enhancers? Nat Genet 47, 8-12 (2015).
65. D. Hnisz et al., Super-enhancers in the control of cell identity and disease. Cell 155, 934-947 (2013).
66. L. S. Ludwig et al., Transcriptional States and Chromatin Accessibility Underlying Human Erythropoiesis. Cell Rep 27, 3228-3240 e3227 (2019).
67. B. D. Pope et al., Topologically associating domains are stable units of replication-timing regulation. Nature 515, 402-405 (2014)

CLAIMS

It is claimed:

1. Use of an agent that increases RNA polymerase II pausing mediated by a mediator complex subunit 26 (MED26) polypeptide in the manufacture of a medicament for promoting erythroid differentiation in an erythroid precursor cell.
2. The use of claim 1, wherein the agent increases the interaction of the MED26 with a transcription pausing factor, such as a factor in a negative elongation factor (NELF), DRB Sensitivity Inducing Factor (DSIF) or polymerase associated factor (PAF) complex, particularly a PAF1, in the cell.
3. The use of claim 1, wherein the agent increases occupancy of a super-enhancer site by the MED26, such as a super-enhancer for a gene selected from CDK6, BCL2, HMGA1, MYB, RPS14, RPL36AL, and RPL27.
4. The use of any one of the foregoing claims, wherein the agent comprises the MED26 polypeptide or an active fragment thereof, or a nucleic acid encoding the active fragment.
5. The use of claim 4, wherein the active fragment comprises an intrinsically disordered region (IDR) of the MED26 polypeptide, more particularly, the active fragment comprises a polypeptide that is at least 90%, such as at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%, identical to the amino acid sequence of SEQ ID NO: 17.
6. The use of claim 5, wherein the active fragment comprises a TFIIS domain and an intrinsically disordered region (IDR) of the MED26 polypeptide, more particularly, the active fragment comprises a polypeptide that is at least 90%, such as at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%, identical to the amino acid sequence of SEQ ID NO: 2.
7. The use of any one of the foregoing claims, wherein the agent comprises the MED26 polypeptide or the nucleic acid encoding the MED26 polypeptide, more particularly, the MED26 polypeptide comprises a polypeptide that is at least 90%, such as at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the amino acid sequence of SEQ ID NO: 1.

8. Use of a polypeptide comprising the amino acid sequence of SEQ ID NO: 17 or a polynucleotide encoding the polypeptide in the manufacture of an agent for promoting erythroid differentiation in an erythroid precursor.
9. The use of claim 8, wherein the polypeptide comprises the amino acid sequence of SEQ ID NO: 2.
10. The use of claim 9, wherein the polypeptide comprises the amino acid sequence of SEQ ID NO: 1.
11. The use of any one of claims 1-10, wherein the erythroid precursor is a hematopoietic stem and progenitor cell (HSPC), such as a CD34+ HSPC.
12. The use of any one of claims 1-10, wherein the erythroid precursor is an erythroblast, such as an erythroblast derived from a CD34+ HSPC.
13. The use of any of claims 1-12, wherein the agent is to be administered to the erythroid precursor *in vitro*, or to the erythroid precursor *in vivo* in a subject in need thereof.
14. The use of claim 13, wherein the subject is in need of treating a disease associated with defective erythroid differentiation.
15. The use of claim 14, wherein the disease associated with defective erythroid differentiation is a myelodysplastic syndrome, such as refractory anemia or refractory cytopenia, erythroid dysplasia, bone marrow failure or megaloblastic anemia.
16. A method of identifying an agent for promoting erythroid differentiation, comprising
    a. providing a composition comprising a MED26 polypeptide or an active fragment thereof;
    b. contacting the composition with a test agent; and
    c. assessing whether the test agent increases the capacity of the MED26 to mediate RNA polymerase II pausing.
17. The method of claim 16, wherein the assessing step comprises assessing whether the test agent promotes the capacity of the MED26 to form an aggregate, to interact with a transcription pausing factor, and/or to occupy a super-enhancer site.
18. Use of a MED26 polypeptide or active fragment thereof in the manufacture of a medicament for delivering a reagent to a cell, wherein the delivering comprises:
    a. preparing a condensate comprising the MED26 polypeptide or active fragment thereof and the reagent; and

b. contacting the cell with the condensate,

whereby the MED26 polypeptide or active fragment thereof delivers the reagent to the cell.

19. The use of claim 18, wherein the reagent is a nucleic acid, such as one or more RNA or DNA molecules, the RNA is such as an mRNA, an siRNA, an antisense RNA, a linear RNA, a circular RNA or a tRNA, the DNA is such as a circular DNA, e.g., a plasmid DNA, or a linear DNA, e.g., an antisense DNA.
20. The use of claim 18 or 19, wherein the MED26 polypeptide or active fragment thereof comprises an amino acid sequence having at least 75%, such as at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% , 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%, sequence identity to SEQ ID NO: 16.
21. The use of claim 20, wherein the MED26 polypeptide or active fragment thereof comprises an amino acid sequence having at least 75%, such as at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% , 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%, sequence identity to SEQ ID NOs: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15.
22. The use of claim 20, wherein the MED26 polypeptide or active fragment thereof comprises the amino acid sequence of SEQ ID NOs: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16.
23. The use of claim 20, wherein the MED26 polypeptide or active fragment thereof comprises the amino acid sequence of SEQ ID NO: 12.
24. The use of claim 20, wherein the MED26 polypeptide or active fragment thereof comprises the amino acid sequence of SEQ ID NO: 2.
25. The use of any one of claim 19, wherein the nucleic acid encodes one or more components for gene editing.
26. The use of claim 25, wherein the gene editing comprises a CRISPR gene editing, and the nucleic acid encodes one or more of a CRISPR RNA (crRNA), a tracrRNA that hybridizes with the crRNA, and a Cas endonuclease, such as Cas9, Cas12, Cas13, preferably, the nucleic acid encodes a single guide RNA comprising the crRNA and the tracrRNA and/or the Cas endonuclease, more preferably, the nucleic acid encodes the single guide RNA and a Cas9 endonuclease.

27. The use of any one of claims 18-26, wherein the step of preparing the condensate comprises mixing the reagent with the MED26 polypeptide or active fragment thereof in a buffer comprising water, wherein the buffer further comprises polyethylene glycol (PEG), preferably the buffer comprises 1-30% (w/v) PEG, such as 2-20% (w/v) PEG, e.g., 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14% or 15% (w/v) PEG, preferably 10% (w/v) PEG.
28. The use of any one of claims 18- 27, wherein the step of preparing the condensate comprises mixing the reagent with 1-200 μM, preferably 5-150 μM, such as 5, 10, 25, 50, 75, 100, 125, or 150 μM, more preferably 10 μM, of the MED26 polypeptide or active fragment thereof.