Genome-Scale CRISPR Screens Identify Novel Genes That Regulate Erythropoietin Production and Secretion

by

Zesen Lin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Pharmacology)
in the University of Michigan
2024

Doctoral Committee:

Associate Professor Rami Naim Khoriaty, Co-Chair
Professor James Alan Shayman, Co-Chair
Professor David Ginsburg
Professor Manojkumar Puthenveedu

Zesen Lin

zesenlin@umich.edu

ORCID iD: 0000-0002-1779-1967

Acknowledgements

I would like to take this opportunity to express my profound gratitude and deep regard to my dissertation advisor, Dr. Rami Khoriaty, for his exemplary guidance and constant encouragement throughout the course of this dissertation. The door to Dr. Khoriaty's office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed me to plan my own work, but steered me in the right direction whenever he thought I needed it.

I would like to express my deepest gratitude to my thesis committee members, Dr. Shayman, Dr. Ginsburg, and Dr. Puthenveedu for their unwavering support and invaluable mentorship throughout the journey of my research. To my committee chair, Dr. Khoriaty, whose expertise and insightful perspectives have been a guiding light in shaping the direction of my work, I am profoundly thankful. I am equally indebted to Dr. Shayman, Dr. Ginsburg and Dr. Puthenveedu, whose critical feedback and encouragement have been instrumental in refining my experiments and methodologies. Their collective wisdom and constructive critiques have not only enriched my academic experience but have also fostered my growth as a scholar. Their dedication to my success and their willingness to invest time and effort into my project have been nothing short of inspiring. I am honored to have worked with such a distinguished and supportive committee, and I hope this work reflects the high standards they espouse.

My sincere thanks also goes to the faculty members of the Pharmacology Department at the University of Michigan for their insightful comments and suggestions, and for challenging me to widen my research from various perspectives.

I am grateful to all my peers in the program for providing a stimulating and fun environment for growth. In particular, I would like to thank Beth McGee, Ginette Balbin-Cuesta, Greggory Meyers, Ann Friedman, Lei Yu, Claire Drysdale, Claire Kerpet, Masaki Ito, and Rillie Saba for their friendship, peer support, and for making the long hours of work more enjoyable.

I am also immensely grateful to my family and friends, who have provided me with moral support and the sense of balance needed to complete this academic work. Special thanks to my parents, Yueping Chen and Jianqiang Lin, whose love and encouragement have been a constant source of strength.

I cannot begin to express my gratitude and appreciation for all these people and their contributions to my academic journey.

Table of Contents

List of Figures

Abstract

Erythropoietin (EPO) is a plasma glycoprotein that binds to bone marrow erythroid progenitors, stimulating their proliferation and differentiation. In adult humans, EPO is secreted into the circulation by hepatocytes and specialized kidney peritubular fibroblasts, contributing approximately 20% and 80% to the total plasma EPO levels, respectively. EPO production is induced by hypoxia due to increased stability of the transcription factor hypoxia-inducible factor (HIF). However, HIF leads to increased expression of genes that promote angiogenesis and oncogenesis, which raises concerns about clinical strategies that promote EPO production via increased HIF levels. Therefore, we seek to identify novel HIF-independent regulators of EPO production as well as regulators of EPO secretion.

To identify novel regulators of EPO secretion, we generated a reporter HEK293T cell line stably expressing EPO fused to GFP, and as an internal control, alpha-1-antitrypsin (A1AT) fused to mCherry. We demonstrated that both EPO and A1AT are efficiently secreted from the reporter cell line, and that treatment with Brefeldin A (which disrupts ER-to-Golgi transport) results in the intracellular accumulation of both proteins, verifying the integrity of the secretion machinery in these cells. The reporter cell line was mutagenized using a CRISPR/Cas9 knockout library, which delivers SpCas9, a puromycin resistance cassette, and a pooled collection of 123,411 single guide RNAs targeting almost every coding gene in the human genome. We isolated cells with normal mCherry but increased or decreased GFP fluorescence. This strategy enabled the

identification of genes that specifically affect EPO but not A1AT levels, thereby ruling out genes that affect global secretion. We identified and validated SURF4 as an ER cargo receptor that is required for the efficient secretion of EPO. Deletion of *SURF4* leads to EPO accumulation within the ER, while over-expression of SURF4 results in enhanced efficiency of EPO secretion.

To identify novel regulators of EPO production, we generated a reporter HEP3B cell line with homozygous insertion of coding sequences for p2A and eGFP at the endogenous *EPO* locus. This HEP3B EPO-p2A-eGFP cell line expresses equivalent levels of EPO and eGFP proteins, both translated from the same mRNA molecule. Using a similar strategy as above, we performed a genome-scale CRISPR knock-out screen with and without HIF activation to identify genes that regulate EPO production independent of HIF. This screen was followed by a secondary validation screen targeting the top 1,255 candidate regulators of EPO production nominated by the genome-scale screen. These studies identified ZNF574 as a novel HIF-independent regulator for EPO production. In normoxia, deletion of *ZNF574* resulted in increased *EPO* mRNA and protein levels. Under hypoxic conditions, deletion of *ZNF574* led to increased EPO production, more so than in wildtype cells exposed to hypoxia. RNA sequencing analysis showed that *ZNF574* deletion did not significantly impact the expression of most HIF-regulated genes. These findings suggest that ZNF574 regulates EPO production without activating the HIF pathway.
In summary, we have identified and validated novel regulators of EPO production and secretion. These findings have important implications for several anemia disorders, particularly for anemia of chronic kidney disease, which results from impaired EPO production.

Chapter 1 Review of Erythropoietin (EPO): Synthesis, Mechanisms, and Regulatory Pathways

1.1 EPO regulates red blood cell (RBC) production.

EPO is a plasma glycoprotein that regulates RBC production. Circulating EPO binds to its cognate cell surface receptor (EPOR) expressed on erythroid progenitors, promoting cell survival, proliferation, and differentiation[1-8]. Consistent with a critical role of EPO in erythropoiesis, mice with biallelic germline deletion of *Epo* (or *EpoR*) exhibit mid-embryonic lethality (at ~E12.5) due to severe anemia[9, 10]. Though EPO has been suggested to have non-hematopoietic functions[11-16], the chief function of EPO is supporting erythropoiesis, since the EPO/EPOR axis is dispensable in non-erythroid cells[2].

1.2 Erythroid stages that depend on EPO.

Multipotent hematopoietic stem cells undergo a series of cell divisions and differentiation steps to generate cells with successively more restricted developmental potential. The first erythroid-committed progenitor termed 'Burst-forming unit-erythroid' (BFU-E) generates "colony-forming units-erythroid" (CFU-E), which in turn generates the proerythroblast. Proerythroblasts differentiate to sequentially generate basophilic, polychromatic, and orthochromatic erythroblasts; the latter cells eliminate nuclei and most organelles as they develop into reticulocytes, which in turn mature to become RBCs. Though EPO may steer hematopoietic progenitors towards the

erythroid lineage[17, 18], it is not required for commitment of bi-/multi-potent progenitors to the erythroid lineage[3]. Indeed, generation of BFU-E and CFU-E does not depend on EPO[3]. In contrast, CFU-Es and proerythroblasts depend on the availability of EPO for their survival and differentiation[1-8].

1.3 EPO structure.

Erythropoietin (EPO) is a glycoprotein hormone with a molecular weight of 30.4 kDa, comprised of 165 amino acids. Its structure features four antiparallel alpha-helices (A, B, C, and D helices), two beta-sheets, and two intra-chain disulfide bridges (Cys7-Cys161 and Cys29-Cys33), which stabilize the alpha helical content by connecting the A and D helices, and the B and C helices, respectively [19, 20].

EPO contains three N-linked glycosylation sites at N24, N38, and N83, which are thought to enhance the molecule's stability and resistance to proteolytic enzymes and one O-linked glycan at S126, which is thought to contribute to its bioactivity. Glycoengineering of these sites has been utilized to create recombinant human EPO with increased activity and a longer duration of action *in vivo* [21].

1.4 Cell origin of EPO production.

The existence of a circulating factor that regulates RBC production had been suggested in 1906[22]; however, it was not until 1950, when a parabiosis experiment demonstrated the humoral regulation of erythropoiesis[23]. In this experiment, the circulation of a rat exposed to hypoxia

was connected to that of another rat exposed to normoxia. Both rats developed enhanced bone marrow erythrocytosis[23], demonstrating the existence of a blood-circulating factor that regulates RBC production[23, 24]. This finding culminated in cloning of the *EPO* gene in 1985[25-27].

1.4.1 Cell of origin of EPO production in adult mammals

In adults, EPO is produced predominantly by the kidney[28-30], and specifically by kidney fibroblasts located in the corticomedullary region[31-34]. Analysis of adult nephrectomized animals demonstrated that ~80% of the total circulating EPO originates from the kidney[35-38]. Therefore, it is no surprise that patients with end stage renal disease often exhibit anemia due to reduced EPO production. Studies in nephrectomized and partially hepatectomized animals identified the liver as the secondary organ of EPO production, accounting for ~20% of the circulating EPO in adults[39].

1.4.2 Cell of origin of EPO production during fetal development

During fetal development, the kidney is dispensable for EPO production. Very early in embryonic development and for a short period of time, EPO is produced primarily in neural cells to support 'primitive' yolk sac erythropoiesis[40]. However, throughout most of fetal development, the main EPO-producing organ is the liver[41-47]. Though the timing of the liver-to-kidney switch in EPO production may vary among species, evidence supports the liver as the major EPO-producing organ until ~1 month after birth in several species, including humans and mice[35]. Cells in the liver that are responsible for EPO production are hepatocytes (and to a lesser extent, stellate cells)[33, 48, 49].

As discussed above, the primary site EPO production shifts throughout development, from neuroepithelial and neural crest cells supporting primitive erythropoiesis, to fetal liver throughout most fetal development, and to the kidney postnatally. The mechanism by which the primary site of EPO production changes throughout development remains unknown. One hypothesis is that the 'adult environment' may repress EPO production in adult hepatocytes. To address this possibility, pre-switch fetal livers (or post-switch adult livers) were transplanted into adult animals[50]. Animals transplanted with fetal (but not adult) livers exhibited a profound (~7 fold) increase in serum EPO levels, irrespective of presence or absence of host kidneys[50]. These findings exclude an extrinsic effect from the adult environment as the etiology for reduced EPO-production in the adult liver. Furthermore, the postnatal switch in EPO production from mostly liver to mostly kidney is not due to alterations in oxygen tension that occur around birth[33]. Though several factors may impact EPO expression[51-53], the exact mechanism by which EPO production is suppressed in adult hepatocytes remains largely unknown.

1.5 *EPO* gene.

In humans, the *EPO* gene is located on chromosome 7 and is composed of 5 exons and 4 introns. *EPO* exons display a high degree of conservation in vertebrates. Interestingly, in contrast to most *EPO* introns, which don't exhibit a high degree of conservation throughout evolution, a portion of *EPO* intron 1 is highly conserved in vertebrates.

1.6 *EPO* enhancer.

Studies using transgenic mouse models suggest that the expression of *EPO* is mediated by distinct enhancers in liver and kidney [45, 54]. In murine renal EPO producing cells, the *EPO*

enhancer is located within a sequence 17.4 kb to 3.6 kb upstream of the *EPO* transcription start site [55]. Although the exact sequence is still unknown, analysis of human *EPO* gene expression in transgenic mice revealed that the renal *EPO* enhancer is located 15 kb to 6kb upstream of the gene [54, 56]. In contrast, the enhancer that regulates hepatic *EPO* expression is located downstream of the gene. Genetic excision of the hepatic *EPO* enhancer results in severe anemia in mice during late embryonic development and the neonatal period, without causing a significant erythropoietic defect in adult mice [45].

1.7 Regulation of EPO by HIF

EPO production is hypoxia-inducible in the liver and kidneys,[28, 30, 35, 39, 50, 57-64] a process that is mediated by the hypoxia-inducible transcription factor, HIF. In the EPO-producing cells of the liver[45, 65] and kidneys[54, 55], HIF binds hypoxia responsive elements in the *EPO* locus, promoting *EPO* transcription[65-67]. HIF is a heterodimer composed of an α (HIF1α or 2α) and a β (HIF1β) subunit. In both liver and kidneys, HIF2α is the main α subunit that controls *EPO* transcription, with a lesser contribution by HIF1α[68].

The stability of the α subunit of HIF, and therefore the level of EPO production, depends on O2 availability. In normoxia, specific HIFα proline residues undergo hydroxylation by prolyl hydroxylase domain (PHD) enzymes (PHD1, PHD2, or PHD3), in a reaction that requires O2. Hydroxylated HIFα is recognized by von Hippel-Lindau (VHL), a component of an E3 ubiquitin ligase complex[69-71], resulting in HIFα polyubiquitination and proteasomal degradation. In addition to HIFα degradation, the binding of HIFα to the transcriptional co-activator p300 is

impaired during normoxia, due to hydroxylation of a specific HIFα asparagine residue by the enzyme 'factor inhibiting HIF' (FIH), a reaction that also requires O2 [71, 72]. Therefore, during normoxia, *EPO* transcription is suppressed, because HIF protein is degraded and because HIF is unable to bind its transcriptional co-activator p300. In contrast, in hypoxia, the PHD and FIH enzymes are inactive; therefore, HIF is not degraded and is able to bind the transcriptional co-activator p300, resulting in transcription of HIF-target genes, such as *EPO*.

1.8 Transcription factors that regulate EPO expression.

As discussed below, several transcription factors, other than HIF, have been shown to regulate *EPO* expression. It is possible that orchestration of several transcription factors may plays a role in the temporal control and tissue-specific expression of *EPO*.

GATA family members. Several members of the GATA family (GATA1,2,3,4 and 6) have been shown to bind a motif (TGCAGATAACAGCCC) in the *EPO* promoter[51, 73, 74]. While binding of GATA4 and GATA6 to the *EPO* promoter leads to enhanced *EPO* expression, binding of GATA1, GATA2 and GATA3 leads to suppressed EPO synthesis [73]. Studies in mice demonstrated that mutations within the GATA binding motif result in aberrant and persistent *EPO* expression in renal distal tubules, renal collecting ducts, hepatic epithelial cells, lung, and thymus [75]. These results suggest that on balance, GATA factors function to limit aberrant EPO production.

WT1 and SP1. Wilms tumor 1 (*WT1*) and specificity protein 1 (*SP1*) genes encodes proteins that synergistically bind the *EPO* promoter (TC**CCC**CAC**CC**C<u>CACC</u>C, with sequences specific to

WT1 highlighted in bold and those specific to SP1 underlined), resulting in upregulation of *EPO* transcription [76, 77]. Overexpression of WT1 was shown to enhance *EPO* expression in HEPG2 cells and induce *EPO* expression in HEK293T cells, which normally don't produce EPO [76]. Loss of WT1 results in marked reduction in *EPO* mRNA levels in fetal liver [76] and adult brain [78]. The interaction between SP1 and WT1 at the *EPO* promoter is cooperative, suggesting a complementary regulatory mechanism in activating *EPO* transcription [76].

HNF4. Hepatic nuclear factor 4 (HNF4) is a transcription factor that plays a role in the regulation of *EPO* expression, particularly in response to hypoxic conditions. HNF4 appears to bind the hepatic *EPO* enhancer, promoting *EPO* expression. The importance of HNF4 in the hypoxia-mediated induction of *EPO* expression was highlighted by experimental evidence demonstrating that reduced HNF4 expression in HEP3B cells leads to marked attenuation of the induction of EPO production mediated by hypoxia [79].

RXRα and RARα. Retinoid X receptor-alpha (RXRα) and retinoic acid receptor (RARα) form a heterodimer that binds to the DR2 motif (TGACCTCTTGACCC) in the *EPO* hepatic enhancer, leading to enhanced *EPO* expression, independent of oxygen tension [53]. Addition of retinoic acid (RA), the common ligand to both receptors, results in increased *EPO* production in primary fetal liver cells and HEP3B cells [53]. Therefore, it is not surprising that deficiency in vitamin A, the precursor of RA, leads to lower kidney *EPO* mRNA levels [80]. RXRα deficient embryos exhibit a transient reduction in hepatic *EPO* mRNA levels at E10.5, which resolves at E12.5 [53]. Notably, *in vitro* studies demonstrated a competition between HNF4 and RXR-RAR heterodimer for binding the DR2 motif in *EPO* [53]. HNF4 appears to be present in excess of

RXR-RAR in hepatocytes after E12.5, which may explain why the expression of *EPO* is not inducible by RA in hepatocytes harvested from embryos older than E12.5 or from postnatal animals.

1.9 *EPO* transcripts.

Recently, 2 human *EPO* transcripts were found to be expressed from the *EPO* locus: i) a functional in-frame canonical transcript (T1) expressed predominantly in the adult kidney and fetal hepatocytes and ii) an out-of-frame transcript (T2) expressed predominantly in adult hepatocytes[81]. The T2 transcript is expressed from an alternative promoter in intron 1 of *EPO* [81]. Notably, a promoter was identified in *EPO* intron 1 by CAGE-seq[81]. The difference in *EPO* transcript expression pattern between fetal hepatocytes/adult kidney on one hand versus adult hepatocytes on the other hand may point towards changes in regulation of *EPO* expression throughout development.

The mechanism by which the functional T1 transcript is replaced by the non-functional T2 transcript in adult hepatocytes is unknown. Several possible explanations for the switch in expression from T1 in fetal hepatocytes to T2 in adult hepatocytes have been proposed. One possible explanation is that a factor that promotes T1 transcription or that represses T2 transcription is expressed in fetal but not adult hepatocytes. An alternative explanation is that a factor that represses T1 transcription or that promotes T2 transcription is expressed in adult but not fetal hepatocytes.

1.10 *EPO* mutations or dysregulation in humans.

Reduced EPO production can arise from congenital loss-of-function mutations in *EPO* [82] or from end stage renal disease [83], resulting in hereditary or acquired anemia, respectively. These disorders should benefit from strategies aimed at increasing the plasma EPO level. Anemia due to other erythroid disorders (ex. a subset of myelodysplastic syndromes) may also benefit from a similar therapeutic strategy, even if the primary defect is not reduced EPO production, as previously shown [84, 85]. While 2 classes of agents (recombinant human EPO and PHD inhibitors) are clinically available for this use, they have significant limitations, as discussed in the introduction to chapter 3.

Similarly, a pathological increase in EPO production may result from congenital etiologies (e.g.: loss-of-function mutations in *VHL* or in the gene encoding *EGLN1*, gain-of-function mutations in *EPAS1*, mutations in hemoglobin resulting in reduced oxygen release, and bisphosphoglycerate mutase deficiency) [86] or from acquired conditions (e.g. tumors that secrete EPO) [86]. These disorders result in polycythemia (increased RBC production) and should benefit from agents that reduce the plasma EPO level. No such agent is currently clinically available.

Chapter 2 The Endoplasmic Reticulum Cargo Receptor SURF4 Facilitates Efficient Erythropoietin Secretion

This chapter is published in MCB. Lin Z, King R, Tang V, et al. The Endoplasmic Reticulum Cargo Receptor SURF4 Facilitates Efficient Erythropoietin Secretion. Mol Cell Biol. 2020;40(23):e00180-20. Published 2020 Nov 6. doi:10.1128/MCB.00180-20

2.1 Abstract

Erythropoietin (EPO) stimulates erythroid differentiation and maturation. Though the transcriptional regulation of *EPO* has been well studied, the molecular determinants of EPO secretion remain unknown. Here, we generated a HEK293T reporter cell line that provides a quantifiable and selectable readout of intracellular EPO levels and performed a genome-scale CRISPR screen that identified SURF4 as an important mediator of EPO secretion. Targeting *SURF4* with multiple independent single guide RNAs (sgRNAs) resulted in intracellular accumulation and extracellular depletion of EPO. Both of these phenotypes were rescued by expression of *SURF4* cDNA. Additionally, we found that disruption of SURF4 resulted in accumulation of EPO in the endoplasmic reticulum (ER) compartment and that SURF4 and EPO physically interact. Furthermore, SURF4 disruption in Hep3B cells also caused a defect in the secretion of endogenous EPO under conditions mimicking hypoxia, ruling out an artifact of heterologous overexpression. This work demonstrates that SURF4 functions as an ER cargo receptor that mediates the efficient secretion of EPO. Our findings also suggest that modulating

SURF4 may be an effective treatment for disorders of erythropoiesis that are driven by aberrant EPO levels. Finally, we show that SURF4 overexpression results in increased secretion of EPO, suggesting a new strategy for more efficient production of recombinant EPO.

2.2 Introduction

Approximately one-third of the proteins encoded by the mammalian genome are secretory proteins [87, 88]. These proteins traffic from the endoplasmic reticulum (ER) to the Golgi apparatus via coat protein complex II (COPII) vesicles before reaching their final destinations: endosomes, lysosomes, plasma membrane, or extracellular space. COPII vesicles have an inner coat composed of SAR1 and SEC23-SEC24 heterodimers and an outer coat composed of SEC13-SEC31 heterotetramers [89]. Though transmembrane cargo proteins may directly interact with COPII components, the physical barrier created by the ER membrane prevents direct interaction between soluble cargos and the COPII coat. Therefore, soluble cargos either passively flow into COPII vesicles (bulk flow) or are captured in COPII vesicles through recognition by intermediary receptors or adapters (cargo capture) [90].

Support for receptor-mediated cargo capture arose from early electron microscopy studies and in vitro assays of cargo packaging in COPII vesicles, which demonstrated efficient selection and concentration of cargos into COPII vesicles, as well as physical interactions between soluble cargos and COPII components [90-95]. Subsequent studies uncovered LMAN1 as the first ER cargo receptor that mediates ER export of soluble cargos in mammals [96-98]. LMAN1, together with its adapter MCFD2, forms a complex that is required for the efficient secretion of

coagulation factors V and VIII, cathepsins, and alpha1-antitrypsin [98-102]. While a handful of additional interactions between soluble cargos and ER receptors have been described in mammals [103], the extent to which bulk flow and cargo capture contribute to recruitment of proteins in COPII vesicles is unclear. This is primarily due to the fact that ER cargo receptors that are necessary for the efficient secretion of the majority of soluble cargos remain unidentified.

Erythropoietin (EPO) is a glycoprotein that is produced predominantly by specialized kidney peritubular fibroblasts and secreted into the plasma [31, 33]. EPO binds to its receptor expressed on erythroid precursors and promotes cell survival, proliferation, and differentiation [1, 2]. EPO plays a crucial role in the regulation of red blood cell production (erythropoiesis). Failure to make sufficient amounts of EPO, as seen in the setting of chronic kidney disease, results in anemia. In contrast, supraphysiological EPO levels, as seen in the context of EPO-secreting tumors, result in polycythemia. Though the transcriptional regulation of EPO production has been well studied [67, 104, 105], the molecular basis of EPO trafficking remains poorly understood.

In this study, in an effort to identify proteins involved in EPO secretion, we developed a genome-scale CRISPR/Cas9 knockout screen that provides a quantifiable and selectable readout of intracellular EPO levels. This screen, followed by several validation experiments, identified the ER cargo receptor SURF4 as a key mediator of efficient EPO secretion. These findings suggest that modulation of SURF4 in the EPO-producing cells could provide a novel strategy for altering

plasma EPO levels and therefore regulating erythropoiesis. Additionally, this report suggests a novel strategy for improving the efficiency of recombinant EPO production.

2.3 Method

2.3.1 Cell Culture.

HEK293T and HEP3B cells were purchased from ATCC. Flp-In T-REx 293 cells were purchased from Invitrogen. HEK293T and Flp-In T-REx 293 cells were cultured in Dulbecco modified Eagle medium (DMEM) (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (Peak Serum) and 1% penicillin-streptomycin (Gibco). HEP3B cells were cultured in alpha-MEM (Gibco) supplemented with 2 mM L-glutamine (Gibco), 10% heat-inactivated fetal bovine serum (Peak Serum), and 1% penicillin- streptomycin (Gibco). All cells were grown in a humidified 37°C incubator with 5% CO_2.

2.3.2 Generation of the EPO-eGFP A1AT-mCherry reporter cell line.

The CMV-EPO-eGFP-p2A-A1AT- mCherry construct was assembled using the NEBuilder HiFi DNA assembly cloning kit (New England Biolabs [NEB]) using vector sequences derived from PCSK9-eGFP-p2A-A1AT-mCherry (32) and *EPO* cDNA obtained from Dharmacon. This construct expresses EPO linked to eGFP via a linker peptide formed of glycines, prolines, and alanines (GGAPAPAPAPAPAPAPAPG). HEK293T cells were transfected with CMV- EPO-eGFP-p2A-A1AT-mCherry using Fugene HD transfection reagent (Promega). Transfected cells were selected with 350 µg/ml Hygromycin (Invitrogen). Five weeks following Hygromycin selection, single cells were sorted into 96-well plates using a SY-3200 flow cytometer (Sony).

Single cell clones were expanded and analyzed for stable expression of EPO-eGFP and A1AT-mCherry using an LSR Fortessa flow cytometer (BD Bioscience).

2.3.3 Generation of the TPO-eGFP A1AT-mCherry reporter cell line.

The TPO cDNA sequence was amplified from human liver RNA (deidentified tissue sample obtained from the tissue procurement core, University of Michigan, institutional review board [IRB] no. HUM00048303) and the CMV-TPO-eGFP-p2A- A1AT-mCherry construct was generated and transfected into HEK293T cells as described in the para- graph above. The same linker peptide linking TPO and eGFP (GGAPAPAPAPAPAPAPAPG) was used. Single cell clones were sorted, expanded, and analyzed for stable expression of TPO-eGFP and A1AT-mCherry as described above.

2.3.4 Expansion and lentiviral preparation of the pLentiCRISPRv2 library.

The pLentiCRISPRv2 whole- genome CRISPR library was obtained from Addgene (Addgene no. 1000000048, a gift from Feng Zhang [45]), expanded by 16 electroporations (8 for each half library) into Endura electrocompetent cells (Lucigen), and plated on 16 24.5-cm bioassay plates (ThermoFisher Scientific). Following a 12- to 14-h incubation at 37°C, colonies were harvested from agar plates, and pooled plasmids for each half library were isolated separately by Maxipreps using an EndoFree Plasmid Maxi kit (Qiagen). To prepare the pooled lentiviral library, 11.3µg of each half library was cotransfected with 17µg of psPAX2 (Addgene no. 12260, a gift from Didier Trono) and 11.3 µg of pCMV-VSV-G (Addgene no. 8454, a gift from Robert Weinberg [96]) using Lipofectamine LTX with PLUS reagent (ThermoFisher Scientific) into each of six T225 tissue culture flasks (ThermoFisher Scientific) containing HEK293T cells

at ~80 to 90% confluence. The medium was changed 24 h post-transfection, and viral supernatant was collected 12, 24, and 36 h afterwards. Media containing viral supernatant were centrifuged at 500*g* for 5 min, pooled, aliquoted, snap-frozen in liquid nitrogen, and stored at -80°C.

2.3.5 CRISPR/Cas9 loss-of-function genome-wide screens.

For each independent screen, more than 142 million reporter cells were plated in 15-cm tissue culture dishes (Corning) at 30% confluence. Cells were transduced with the lentiviral library (with 8 μg/ml Polybrene [Sigma]) at a multiplicity of infection (MOI) of ~0.3. Twenty-four hours after viral transduction, puromycin selection (1 μg/ml; Sigma) was applied for 4 days. Subsequently, cells were kept at a logarithmic phase of growth and passaged every 2 or 3 days, maintaining more than 36 million cells in culture at all times in order to preserve library depth. Fourteen days post-transduction, ~80 million cells were isolated from tissue culture dishes using 0.25% trypsin (Gibco), pelleted by centrifugation (350 *g*, 4°C, 5 min), resuspended in cold phosphate buffered saline (PBS) plus 2% fetal bovine serum (FBS), and filtered through a 35-μm mesh into flow cytometry tubes (Corning). Cells were divided into 20 tubes and maintained on ice until sorting. Cells with normal mCherry fluorescence (mid 70 to 80% fluorescence) and top or bottom ~7% eGFP fluorescence (~4 million cells/population) were sorted using a BD FACSAria III (BD Biosciences) and collected into 15-ml polypropylene tubes (Cellstar) containing media. Genomic DNA was extracted using a DNeasy blood and tissue kit (Qiagen), and integrated lentiviral sgRNA sequences were amplified by a two-step PCR (20 cycles and 14 cycles, respectively) as previously described [106, 107] using a Herculase II fusion DNA polymerase kit (Agilent Biotechnologies). DNA was purified after each of the PCRs using a

QIAquick PCR purification kit (Qiagen). Following the two-step PCR, DNA was analyzed with a bioanalyzer (Agilent), and the sgRNA amplicons were sequenced using a NextSeq 500 sequencing system (Illumina). On average, 23.5 million reads were generated for each sorted cell population of each screen. Overall, 98% of the reads had a per sequence quality score (phred-based base quality score) of greater than 30. A total of 104,331 sgRNA sequences were mapped and identified (along with the barcode corresponding to each cell population of each replicate) using a custom Perl script as previously described [108]. Enrichment at the sgRNA and gene levels was analyzed using DESeq2 and MAGeCK, respectively [109, 110].

2.3.6 Disruption of candidate genes using CRISPR/Cas9.

sgRNAs targeting several genes and several non-targeting sgRNAs were cloned into the pLentiCRISPRv2 plasmid (Addgene no. 52961, a gift from Feng Zhang [107]) as previously described [111]. pLentiCRISPR plasmids were packaged into lentivirus, using the same method described above. To disrupt genes in a population of cells, cells were transduced with lentivirus at an MOI of ~0.3. Subsequently, transduced cells were selected with puromycin and passaged for 10 to 14 days prior to FACS analysis. For all validation experiments, a minimum of three biologic replicates were analyzed. The non-targeting sgRNA sequences are sgRNA1 (GTTCATTTCCAAGT CCGCTG), sgRNA2 (CGTGTGTGGGTAAACGGAAA), sgRNA3 (GTATTACTGATATTGGTGGG), and sgRNA4 (TCATGCTTGCTTGGGCAAAA). *SURF4*-targeting sgRNA sequences are sgRNA1 (TCAGACAGAGGC GCGCCACG) and sgRNA2 (CAGGTAGCCGCAGTTCCAGG).

2.3.7 Generation of SURF4-deficient clonal cell lines.

To generate clonal cell lines that are deficient for SURF4, a sgRNA targeting *SURF4* exon 2 was cloned into the PX459 plasmid (Addgene no. 62988, a gift from Feng Zhang) as previously described [112], and the construct was transiently transfected into cells using Fugene HD transfection reagent (Promega). Twenty-four hours post-transfection, puromycin (1µg/ ml; Sigma) selection was applied for 3 days, and subsequently, single cells were sorted into each well of three 96-well plates using the SY-3200 flow cytometry instrument (Sony). Following expansion of the single cell clones, genomic DNA was extracted with QuickExtract (Epicentre) and indels were determined by amplification of the sgRNA target site by PCR using Herculase II Fusion DNA polymerase (Agilent Biotechnologies) and Sanger sequencing. The primers used for PCR and Sanger sequencing are *SURF4* forward (TCTGTTCCTCACACACCCCGCCC) and *SURF4* reverse (ACTCACTCAGCTGTCCCAGCAAG). Three independent single cell clones with homozygous frameshift indels in *SURF4* (13-bp insertion, 1-bp insertion, and 10-bp deletion) were generated.

2.3.8 Flow cytometry analysis.

HEK293T cells were detached with 0.25% trypsin (Gibco), washed with PBS, collected by centrifugation (350 *g*, 5 min, 4°C), resuspended in cold PBS with 0.1% bovine serum albumin (BSA) and 10 mM HEPES (Invitrogen), filtered with 70-µm cell strainers, and analyzed by BD LSR Fortessa (BD Bioscience). FlowJo (Tree Star) was used to calculate the mean fluorescence intensity and for further analysis.

2.3.9 Brefeldin A treatment.

HEK293 cells stably expressing EPO-eGFP (or TPO-eGFP) and A1At-mCherry were incubated with 1 μg/ml brefeldin A (Biolegend) for 12 h. Subsequently, cells were collected as described above and analyzed by flow cytometry for intracellular accumulation of EPO-eGFP (or TPO-eGFP) and A1AT-mCherry.

2.3.10 Western blots.

To prepare cell lysates, cells were washed in PBS, suspended in radioimmunoprecipitation assay (RIPA) buffer (Invitrogen) supplemented with cOmplete protease inhibitor cocktail (Sigma), briefly sonicated, and incubated for 30 min in the cold room with end-over-end rotation. Cell lysates were cleared by centrifugation to remove cell debris (20,000 g, 30 min, 4°C) and were analyzed immediately or stored at -80°C until analysis. Protein quantification was performed using Pierce BCA protein assay kit (ThermoFisher Scientific) per the manufacturer's instructions. Lysates were boiled for 5 min at 95°C with 4x Laemmli sample buffer (Bio-Rad) supplemented with beta-mercaptoethanol. Equal amounts of proteins were loaded on either a 4 to 12% Bis-Tris gel or a 4 to 20% Tris-glycine gel (Invitrogen), and sodium dodecyl sulfate (SDS) gel electrophoresis was performed as previously described [113]. Proteins were then transferred onto a nitrocellulose membrane (Bio-Rad). Following blocking in 5% (wt/vol) milk-Tris-buffered saline with Tween (TBST), membranes were incubated with primary antibody at 4°C overnight, washed three times in TBST, probed with peroxidase-coupled secondary antibodies, washed again three times in TBST, and developed with SuperSignal West Pico Plus (ThermoFisher Scientific). For horseradish peroxidase (HRP)-conjugated primary antibodies, nitrocellulose membranes were incubated with these antibodies and immediately developed following three TBST washes. Densitometry was performed with ImageJ. To test for the

secretion efficiency of various cargo proteins, cells were seeded at equal densities in six-well plates or 10-cm plates, and conditioned medium was collected at different time points, cleared by centrifugation (500 g, 5 min, 4°C), and analyzed immediately (by Western blotting) as described above or stored at -80°C until analysis.

2.3.11 Antibodies.

The following antibodies were used for immunoblotting: anti-GFP (Abcam, ab290), anti-mCherry (Abcam, ab167453), anti-calnexin (Cell Signaling, 2679S), anti-glyceraldehyde-3-phosphate dehydrogenase (anti-GAPDH) (Millipore, MAB374), HRP-conjugated anti-FLAG (Abcam, ab1238), anti- alpha-tubulin (Abcam, ab176560), HRP-conjugated anti-mouse IgG (Bio-Rad, catalog no. 1706516), and HRP-conjugated anti-rabbit IgG (Jackson ImmunoResearch Laboratories, catalog no. 111-035-003).

2.3.12 Tetracycline-induced EPO-FLAG expression.

The coding sequence of EPO with a C-terminal FLAG was cloned into pDEST-pcDNA5-BirA-FLAG (Invitrogen) using NEBuilder HiFi DNA assembly cloning kit (NEB). Wild-type, SURF4-deficient (with homozygous frameshift *SURF4* indels), or SURF4-rescued (with homozygous frameshift *SURF4* indels but with stable expression of *SURF4* cDNA) Flp-In T-REx HEK293 cells with tetracycline-inducible expression of EPO-FLAG were generated as previously described [114]. The canonical isoform of *SURF4* encoded by NM_033161.3 was used. To induce the expression of EPO-FLAG, tetracycline (1 μg/ml) was added to the medium. Cells and medium were collected prior to the addition of tetracycline and 12 and 24 h following

tetracycline. Intracellular and extracellular EPO levels were analyzed by Western blotting as described above.

2.3.13 EndoH assay.

HEK293T cells that are either wild type, SURF4 deficient, or SURF4 rescue (defined in the paragraph above) were transfected with a plasmid expressing EPO-eGFP. Thirty-six hours post-transfection, total cell lysates were prepared, and protein quantification was performed, both as described above. Lysates were incubated with denaturing buffer (NEB) 95°C for 10 min, and equal amounts of lysates (180 µg) were treated with either 1 µl of endoglycosidase H (EndoH) (NEB), peptide-*N*-glycosidase F (PNGase F) (NEB), or dimethyl sulfoxide (DMSO) as the control for 1 h (37°C). Subsequently, Laemmli buffer (Bio-Rad) was added, and the samples were boiled (95°C) for 5 min. Samples were loaded on a 4 to 12% Bis-Tris gel (Invitrogen), and Western blotting was performed as described above. This experiment was performed in biologic triplicates.

2.3.14 Live-cell confocal fluorescence microscopy.

Wild-type or SURF4-deficient HEK293T cells that stably express EPO-eGFP and A1AT-mCherry were transfected with a plasmid expressing ERoxBFP (Addgene no. 68126, a gift from Erik Snapp [115]). Twenty-four hours post-transfection, cells were seeded on Lab-Tek Chambered Coverglass (ThermoFisher). Fluorescent images were captured on a Nikon A2 confocal microscope. To quantify colocalization between two proteins, Pearson correlation coefficient was calculated using the Nikon Elements software. This experiment was performed

and analyzed by an investigator in a blind manner (investigator unaware of the genotype of the cells).

2.3.15 Coimmunoprecipitation.

Flp-In T-Rex 293 cells that are either wild type, SURF4 deficient, or SURF4 rescued (defined above) were transfected with CMV-EPO-eGFP-p2A-A1AT-mCherry using Fugene HD transfection reagent (Promega). Twenty-four hours posttransfection, cells were washed with PBS and incubated in PBS containing 2mM dithiobis (succinimidyl propionate) (Pierce) for 30min at room temperature. Subsequently, 20 mM Tris-HCl (pH 7.5) was added to quench the reaction. Cells were then washed twice in PBS, and cell lysis was performed with the following lysis buffer (100 mM Tris, 10% glycerol, 1% NP-40, 130 mM NaCl, 5 mM $MgCl_2$, 1 mM NaF, and 1 mM EDTA, supplemented with cOmplete protease inhibitor cocktail [pH 7.5]). Cell lysates were collected as described above and incubated overnight at 4°C with either anti-FLAG M2 magnetic beads (Sigma) or GFP-Trap beads (ChromoTek). Following five washes with lysis buffer, proteins were eluted from the beads via incubation with 2x Laemmli sample buffer containing beta-mercaptoethanol for 15 min at room temperature.

2.3.16 Generation of cell lines expressing low or high SURF4 levels.

A construct expressing SURF4 and the Katushka2S fluorescent marker (PGK-SURF4-p2A-Katushka2S) was assembled with the NEBuilder HiFi DNA assembly cloning kit (NEB) using vector sequence derived from LV1-5 (Addgene no. 68411) and cDNAs of *SURF4* and *Katushka2S* (a gift from Gary Luker [116]). The canonical isoform of SURF4 encoded by

NM_033161.3 was used. The construct was packaged into lentivirus as described above and transduced at a MOI of ~1 into Flp-In T-REx 293 or HEP3B cells. Transduced cells were selected with puromycin and passaged for 14 days prior to FACS sorting. Cells with top and bottom 10% Katushka2S fluorescence were sorted.

2.3.17 Generation of SURF4-deficient HEP3B cells.

Wild-type HEP3B cells were transduced with lentiviral sgRNA targeting *SURF4*, control sgRNA (combination of nontargeting sgRNAs and sgRNAs targeting genes that do not affect EPO such as *BCL11A*, *MPL*, and *SERPINA1*), or sgRNA targeting *EPO* as a positive control. Cells were selected with puromycin and passaged for at least 2 weeks prior to further analysis. EPO levels in the conditioned media were compared between *SURF4*-deleted cells and control cells, correcting for the total cell number at the time of EPO measurement. Genomic DNA was extracted from HEP3B cells using QuickExtract (Epicentre).

2.3.18 EPO ELISA.

Equal numbers of cells were seeded in 6-well or 24-well plates. For HEP3B cells, EPO production was stimulated with $CoCl_2$ (75 μM, Sigma) for 24 h and conditioned medium was collected and cleared by centrifugation (500x *g*, 5 min, 4°C). For Flp-In T-Rex HEK293 cells with tetracycline- inducible expression of EPO-FLAG, tetracycline (1 μg/ml) was added for 12 h and conditioned medium was collected, cleared by centrifugation (500x *g*, 5 min, 4°C), and diluted 1:500. EPO level was measured in the conditioned medium using the Legend Max human

erythropoietin enzyme-linked immunosorbent assay (ELISA) kit (Biolegend), according to the manufacturer's instructions.

2.3.19 qRT-PCR.

RNA was isolated using the RNeasy kit (Qiagen), and cDNA was synthesized using the Superscript first-strand synthesis system for reverse transcription-PCR (RT-PCR) (Invitrogen) with random hexamers. Quantitative RT-PCR (qRT-PCR) was perfomed using Power SYBR green PCR Master Mix using the ViiA7 real-time qPCR System from ThermoFisher. Samples were analyzed in triplicates, and *A1AT* (forward primer [GGTCAACTGGGCATCACTAA] and reverse primer [GATGGTCAGCACAGCCTTAT]) was used as an internal control for EPO (forward primer [GGGAGCCCAGAAGGAAGCCAT] and reverse primer [CTGCAGGCCTCCCCTGTGTA]). *TUBA1A* (forward primer [CGATATTGAGCGTCCAACCTAT] and reverse primer [TTCAGGGCTCCATCAAATCTC]) was used as an internal control for *SURF4* (forward primers CTCTTGTTGTGTGGCTCTTTG and GATGAGGAACCTGGCCCTGGG; respective reverse primers TGGTCTGGA AGAAGTCGTATTT and CCTGCCTCCGAGCTGCATGTA). The threshold cycle ($2^{-\Delta\Delta CT}$) method was used to determine relative gene expression.

2.3.20 Statistical analysis.

CRISPR screen data analysis was performed as described above. The statistical differences in mean fluorescence intensity between EPO-eGFP and A1AT-mCherry were compared by two-way analysis of variance (ANOVA). The difference in extracellular EPO-FLAG level among

wild-type, SURF4-deficient, and SURF4-rescued Flp-In T-REx 293 cells were compared by two-way ANOVA. The Pearson correlation coefficient differences between wild-type and SURF4-deficient HEK293T cells were compared by unpaired *t* test. The statistical difference in extracellular EPO detected by EPO ELISA was assessed using an unpaired *t* test. The differences in the relative amounts of EndoH-sensitive EPO among wild-type, SURF4-deficient, and SURF4-rescued HEK 293T cells were assessed by one-way ANOVA.

2.4 Results.

2.4.1 Generation of a reporter cell line that allows for a quantifiable and selectable readout of intracellular EPO levels.

To identify proteins that regulate the intracellular trafficking of EPO, we developed a genome-scale functional screen that provides a quantifiable and selectable readout of intracellular EPO accumulation. Therefore, we generated a reporter HEK293T cell line stably expressing enhanced green fluorescent protein (eGFP)-tagged EPO (EPO-eGFP) and, as an internal control, mCherry-tagged alpha1-antitrypsin (A1AT-mCherry) (Fig. 2-1A). This cell line is herein referred to as the EPO-eGFP/A1AT-mCherry reporter cell line or just as the reporter cell line. Importantly, EPO-eGFP and A1AT-mCherry are equally expressed from the same cytomegalovirus (CMV) promoter, with their respective coding sequences separated by a sequence encoding a P2A peptide (Fig. 2-1A).

We found that both EPO and A1AT are efficiently secreted from the reporter cell line (Fig. 2-1B and C) and that disruption of ER-to-Golgi transport with brefeldin A results in intracellular accumulation of EPO and A1AT (Fig. 2-1D). Deletion of the gene for the ER cargo receptor for

A1AT, *LMAN1*, resulted in intracellular accumulation of A1AT, as expected, with no effect on intracellular EPO (Fig. 2-1E), ruling out a role for LMAN1 in EPO secretion. These studies demonstrate that the machinery required for the efficient secretion of EPO via the classical secretory pathway is intact in this cell line and establish the utility of this cell line to identify modifiers of intracellular EPO levels.

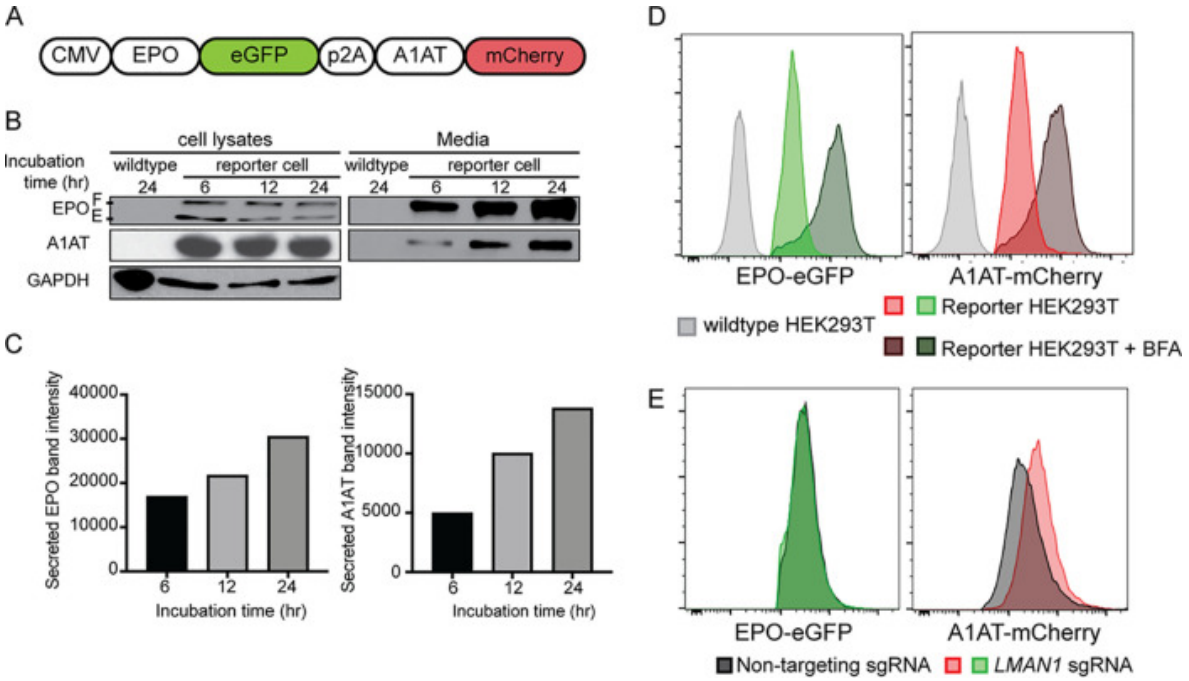

Figure 2–1 A reporter HEK293T cell line stably expressing EPO-eGFP and A1AT-mCherry. (A) A construct that expresses EPO-eGFP and A1AT-mCherry from the same CMV promoter was assembled and used to generate the reporter cell line. A P2A sequence separates EPO-eGFP from A1AT-mCherry. (B) Intracellular and extracellular EPO-eGFP and A1AT-mCherry protein abundance was determined by Western blotting using anti-eGFP and anti-mCherry antibodies, respectively. E, ER form of EPO; F, fully glycosylated EPO. (C) Protein abundance was quantified using ImageJ. (D) Inhibiting ER-to-Golgi transport with brefeldin A (BFA) leads to intracellular accumulation of EPO-eGFP and A1AT-mCherry, as measured by fluorescence intensity. (E) LMAN1 deletion results in intracellular accumulation of A1AT with no effect on EPO. Unless stated otherwise, data are represented as mean values.

2.4.2 A CRISPR/Cas9 loss-of-function screen identified SURF4 as putative regulator of intracellular EPO level.

To identify proteins that affect EPO secretion, we mutagenized the EPO-eGFP/A1AT-mCherry reporter cell line with a CRISPR/Cas9 knockout library (hGeCKO-v2), which delivers SpCas9, a puromycin resistance cassette, and a pooled collection of 123,411 single guide RNAs (sgRNAs) that include 6 sgRNAs targeting nearly every gene in the human genome. Transduction of the library was performed at a low multiplicity of infection (MOI) (~0.3), such that most infected cells receive one sgRNA to mutate one gene in the genome. Puromycin selection was applied from days 1 to 4 post-transduction. After an additional 9 days, cells with normal mCherry but increased (top ~7%) or decreased (bottom ~7%) eGFP fluorescence were isolated (Fig. 2-2A). This cell sorting strategy allows the identification of genes that affect EPO but not A1AT levels, therefore reducing the likelihood of identifying genes that affect global protein secretion. Integrated sgRNA sequences were quantified by deep sequencing and analyzed for their enrichment in the eGFP high compared to the eGFP low population.

This screen, performed in biological triplicates, identified that the sgRNA sequences targeting only one gene, the surfeit locus protein 4 gene (*SURF4*), are enriched in the eGFP high population compared to the eGFP low population at a false discovery rate (FDR) of <10% (Fig. 2-2B). Five out of six sgRNAs targeting *SURF4* were significantly enriched in the eGFP high population (Fig. 2-2C and D).

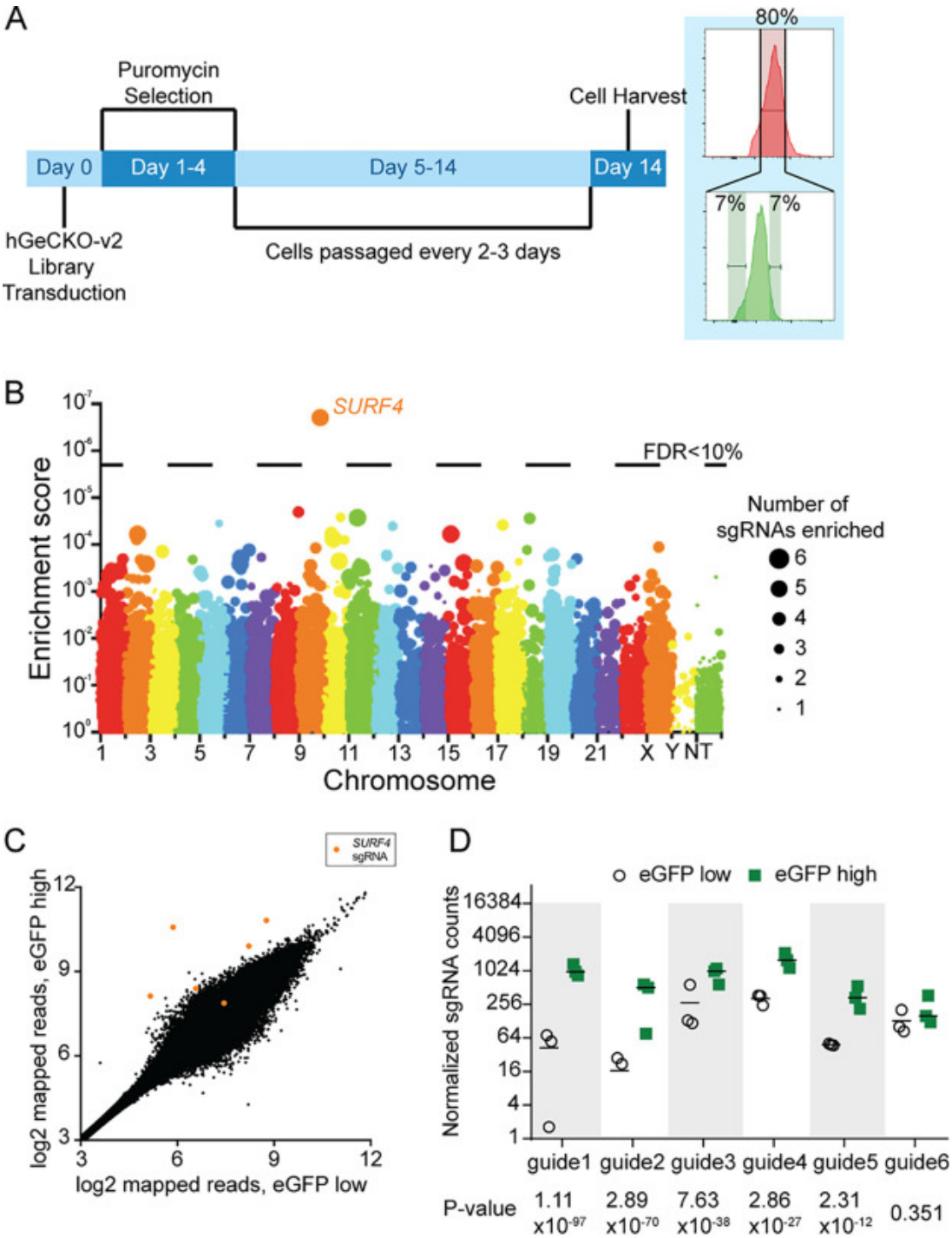

Figure 2–2 CRISPR/Cas9 loss-of-function screen to identify genes that affect intracellular EPO levels. (A) Screen strategy. Twenty-four hours following transduction of the CRISPR library, puromycin selection was applied for 3 days. At day 14, cells with unchanged mCherry but with top or bottom 7% eGFP fluorescence were isolated. sgRNA abundance was then determined in each cell population. (B) Gene level enrichment score was calculated for every gene using MAGeCK (see Materials and Methods). Each gene is represented by a bubble, the size of which is proportional to the number of sgRNAs with significant enrichment in the eGFP high population. *SURF4* has the highest MAGeCK enrichment score and is the only gene for which the false discovery rate (FDR) is <10%. NT, nontargeting. (C) Normalized abundance of SURF4-targeting sgRNAs in the eGFP high and eGFP low populations.

The abundance score was calculated from three biological replicates, using DEseq (see Materials and Methods). SURF4-targeting sgRNAs are highlighted in orange. (D) Normalized counts for the six SURF4-targeting sgRNAs included in the library, for all three biological replicates. P values were calculated using MAGeCK. Unless stated otherwise, data are represented as mean values and the error bars represent standard deviations.

2.4.3 SURF4 deletion results in intracellular accumulation and reduced secretion of EPO.

To validate the results of the screen, we targeted *SURF4* with one sgRNA (sgRNA1) that showed significant enrichment in the whole-genome screen (Fig. 2-2D) and a second sgRNA (sgRNA2) not included in the hGeCKO-v2 library. *SURF4* mutagenesis with sgRNA1 or sgRNA2 was highly efficient, resulting in insertions or deletions (indels) in ~97% and 77% of alleles, respectively (Fig. 2-3A). Cells transduced with *SURF4* sgRNA1 or sgRNA2 exhibited increased intracellular accumulation of EPO-eGFP, with no effect on A1AT-mCherry (Fig. 2-3B and C). This finding was confirmed in three independent EPO-eGFP/A1AT-mCherry reporter cell clones (Fig. 2-3D), ruling out an artifact unique to the clone used in the original screen.

To further confirm a direct effect of SURF4 deficiency on intracellular EPO accumulation, we next generated three clonal reporter cell lines with confirmed frameshift mutations of both *SURF4* alleles by transient expression of *SURF4* sgRNA1. The increased intracellular EPO protein levels observed in *SURF4*-deleted cells was completely rescued by a lentivirus expressing wild-type *SURF4* cDNA (Fig. 2-3E and F), ruling out an off-target effect shared by sgRNA1 and sgRNA2. Taken together, these findings demonstrate that *SURF4* disruption results in intracellular accumulation of EPO. Notably, disruption of ER-to-Golgi transport with brefeldin A further enhances the intracellular accumulation of EPO-eGFP in *SURF4*-deleted cells (Fig. 2-3G).

A

sgRNA	*SURF4* sgRNA 1	*SURF4* sgRNA 2
Number of alleles with indels/total number of alleles tested	37/38	17/22
Number of alleles with frameshift mutation/total number of alleles tested	29/38	13/22

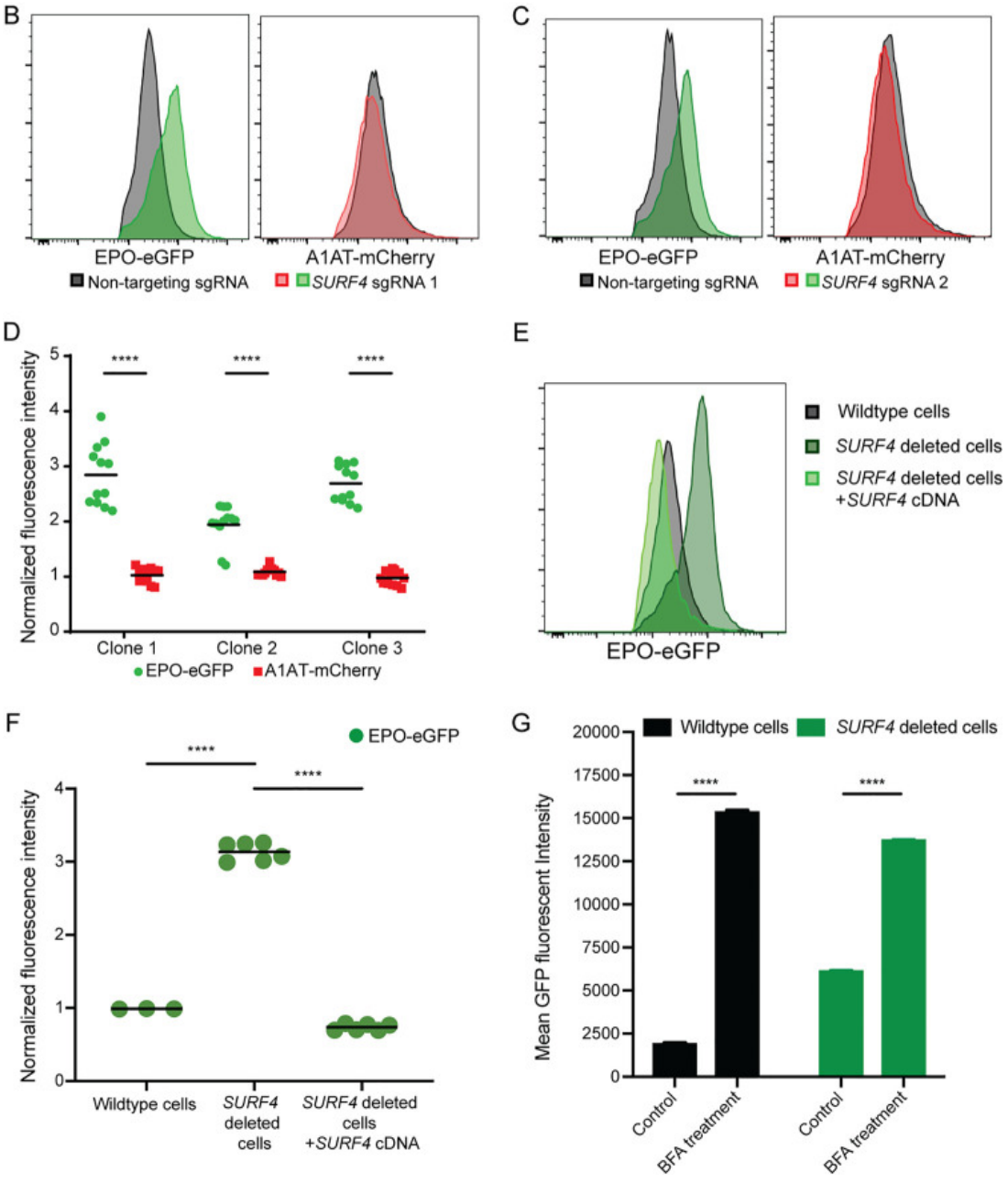

Figure 2–3 SURF4 deletion results in intracellular accumulation of EPO-eGFP. (A) SURF4-targeting sgRNA1 and sgRNA2 are highly efficient, causing indels in ~ 97% and 77% of alleles, respectively. (B and C) Flow cytometry histograms showing intracellular accumulation of EPO, but not A1AT, following SURF4 deletion in HEK293T cells, using two independent sgRNAs, sgRNA1 (B) or sgRNA2 (C). (D) Quantification of intracellular mean fluorescence intensity in three independent clonal reporter cell lines transduced with SURF4-sgRNA1 (n = 12). Results were normalized to mean fluorescence intensity of cells transduced with nontargeting sgRNAs. (E and F) Flow cytometry histograms and normalized mean fluorescence intensity of EPO-eGFP in several clonal cell lines

with sequence-confirmed SURF4 frameshift mutations (SURF4 deleted) with or without stable expression of wild-type SURF4 cDNA. Mean fluorescence intensity in panel F was normalized to that of wild-type cells. (G) Mean eGFP fluorescence intensity in wild-type and SURF4-deleted reporter cells with and without treatment with BFA (n = 3 per condition). ****, $P < 0.0001$.

To rule out an indirect effect on EPO-eGFP secretion resulting from an interaction between eGFP and SURF4, we analyzed the dependence of FLAG-tagged EPO on SURF4 for secretion. We generated a wild-type and SURF4-deficient HEK293 cell line expressing FLAG-tagged EPO (EPO-FLAG) from a tetracycline-inducible promoter (Fig. 2-4A). Following tetracycline administration, the intracellular EPO accumulation was significantly more pronounced in SURF4-deficient cells than in wild-type cells (Fig. 2-4B), recapitulating the findings described above with EPO-eGFP and ruling out an indirect effect due to the eGFP tag.

SURF4 localizes to the ER membrane [106, 117, 118] and functions as an ER cargo receptor, suggesting that the increased accumulation of intracellular EPO in the setting of SURF4 deficiency is secondary to reduced EPO secretion. Consistent with this hypothesis, the extracellular EPO-FLAG protein level was considerably lower in the conditioned media of SURF4-deleted cells compared to the conditioned media of wild-type cells (Fig. 2-4B and C), as was the ratio of extracellular to intracellular EPO-FLAG levels (Fig. 2-4D). The latter findings observed in SURF4-deficient cells were rescued by stable expression of *SURF4* cDNA (Fig. 2-4B and C). These results indicate that disruption of *SURF4* results in a defect in EPO secretion.

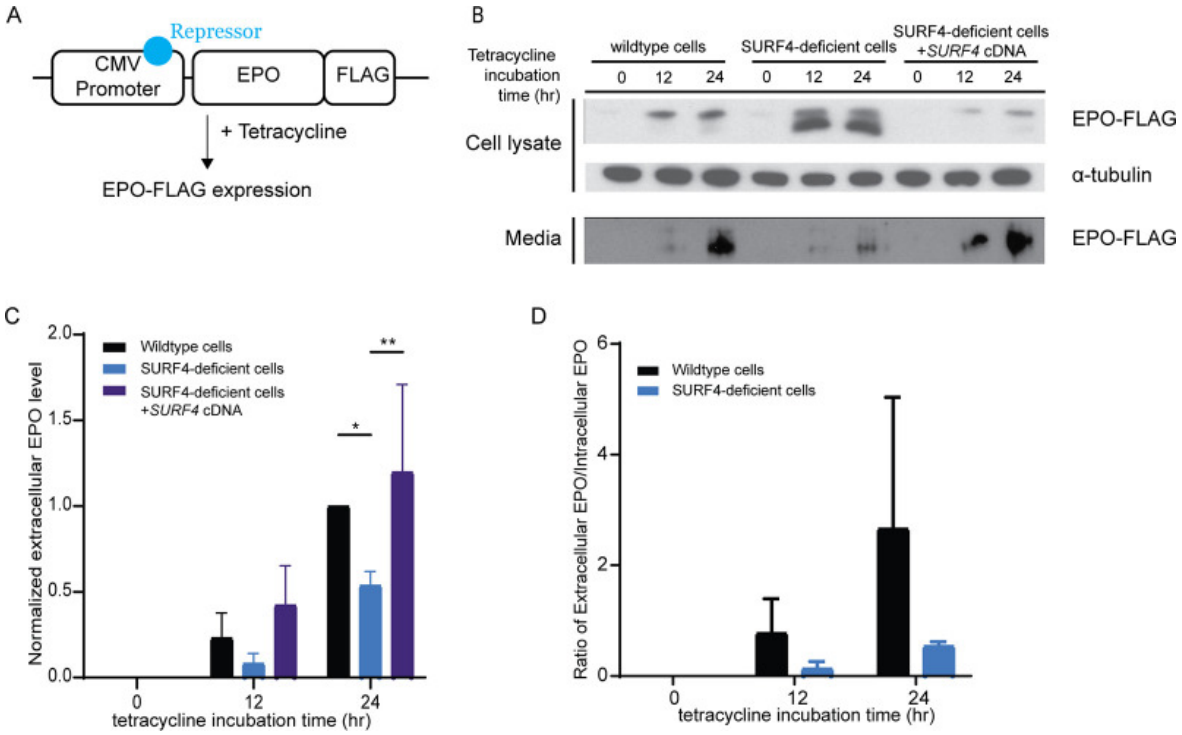

Figure 2–4 SURF4 mutagenesis causes reduced extracellular EPO-FLAG secretion. (A) We generated a Flp-In TREX HEK293 cell line with tetracycline-inducible EPO-FLAG expression. (B) Intracellular and extracellular EPO-FLAG abundance in wild-type, SURF4-deficient, and SURF4-rescued cells was measured by Western blotting (using anti-FLAG antibody) after 0, 12, and 24 h of incubation with tetracycline. α-Tubulin was used as a loading control. (C and D) Quantification of densitometry of extracellular EPO (C) and ratios of extracellular/intracellular EPO normalized to α-tubulin in 3 independent experiments (D). *, $P < 0.05$; **, $P < 0.01$ by two-way ANOVA.

2.4.4 SURF4 deletion results in accumulation of EPO in the ER.

We next performed live-cell fluorescent confocal microscopy to determine the localization of accumulated EPO in the setting of *SURF4* deletion. We cotransfected the EPO-eGFP/A1AT-mCherry reporter construct (Fig. 2-1A) with a vector expressing an ER blue fluorescent marker (ERoxBFP) into wild-type or SURF4-deficient HEK293 cells. We quantified the degree of colocalization between EPO and ERoxBFP (as well as A1AT and ERoxBFP, as a control) by Pearson correlation coefficient (PCC). SURF4-deficient cells exhibited an increased colocalization of EPO (but not A1AT) with ERoxBFP compared to wild-type cells (PCC of 0.7870 in *SURF4*-deleted cells versus 0.2934 in wild-type cells, $P < 0.0001$) (Fig. 2-5A and B).

To confirm the ER accumulation of EPO upon *SURF4* disruption, we tested the glycosylation status of EPO in SURF4-deficient cells. EPO contains three N-glycosylation sites [119-121]. In the ER, N-linked high-mannose oligosaccharides are added to EPO and further modifications are made in the Golgi apparatus. The ER form of EPO is cleavable by endoglycosidase H (EndoH) [122, 123], but the post-ER form is not. Therefore, the ratio of EndoH-cleaved to EndoH-uncleaved EPO will approximate the ratio of the amount of EPO in the ER versus the amount of EPO in the Golgi apparatus or beyond. In SURF4-deficient cells, the ratio of ER/post-ER form of EPO was significantly increased compared to that in wild-type cells (Fig. 2-5C to E), an effect that was decreased by stable expression of *SURF4* cDNA (Fig. 2-5C and E). Taken together, these results demonstrate that SURF4 promotes the efficient ER exit and secretion of EPO.

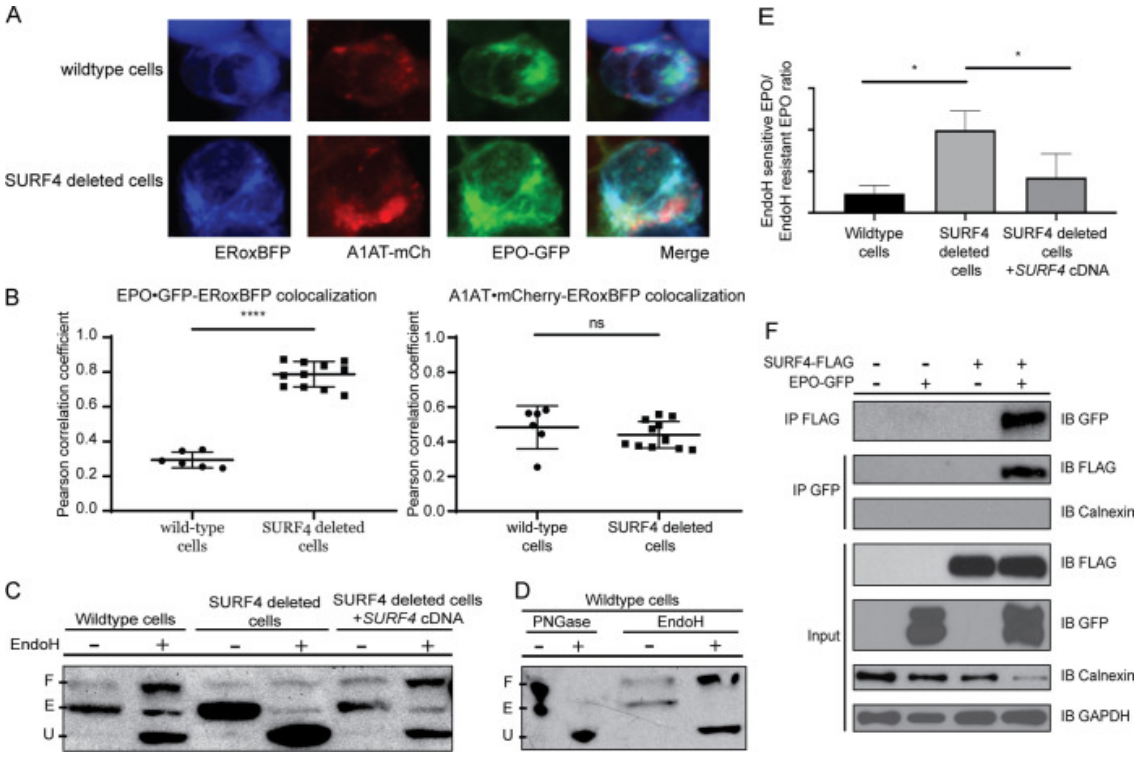

Figure 2–5 Disruption of SURF4 results in accumulation of EPO in the ER. (A) Live-cell fluorescent confocal microscopy of wild-type or SURF4-deleted reporter cells expressing the ER marker, ERoxBFP. (B) Quantification of the degree of colocalization between EPO and ERoxBFP, as well as A1AT and ERoxBFP as a control, by Pearson correlation coefficient. n = 6 for wild-type, n = 11 for SURF4-deficient cells. ****, $P < 0.0001$ by unpaired Student's t test; ns, not significant. (C) Cell lysates were collected from wild-type, SURF4-deleted, or SURF4-rescued cells (SURF4-deleted cells with stable expression of wild-type SURF4 cDNA) expressing EPO-eGFP and were either treated with EndoH or left untreated. Immunoblotting was done with anti-eGFP antibody. The different forms of EPO (E, ER form of EPO [EndoH sensitive]; U, unglycosylated EPO; F, fully glycosylated EPO [post-Golgi form of EPO]) were confirmed by treating wild-type cells with either PNGase or EndoH (D). (E) Quantification of EndoH sensitivity from three independent experiments. *, $P < 0.05$. (F) FLAG antibody or eGFP antibody was used to immunoprecipitate EPO-eGFP or SURF4-FLAG, respectively, from lysates of cells expressing either EPO-eGFP, SURF4-FLAG, both, or neither. IP, immunoprecipitation; IB, immunoblotting.

2.4.5 SURF4 physically interacts with EPO.

To determine whether SURF4 binds to EPO, we tested for reciprocal coimmunoprecipitation of SURF4-FLAG and EPO-eGFP in HEK293T cells. An antibody against the FLAG epitope coimmunoprecipitated EPO-eGFP but not the ER luminal resident protein calnexin. Similarly, an antibody against eGFP coimmunoprecipitated FLAG-SURF4 (Fig. 2-5F). These results are consistent with a specific physical interaction between SURF4 and EPO.

Thrombopoietin (TPO) shares significant sequence homology with EPO. To test whether TPO, similarly to EPO, depends on SURF4 for efficient secretion, we generated two independent clonal HEK293 cells stably expressing and efficiently secreting TPO-eGFP and A1AT-mCherry (Fig. 2-6A and B). As expected, disruption of ER-to-Golgi transport with brefeldin A results in intracellular accumulation of TPO and A1AT (Fig. 2-6C). Notably, like A1AT, TPO did not accumulate intracellularly upon *SURF4* deletion (Fig. 2-6D and E). These findings demonstrate the specificity of SURF4 for promoting EPO secretion and suggest that the SURF4-EPO interaction is mediated by one of the EPO domains not present in TPO.

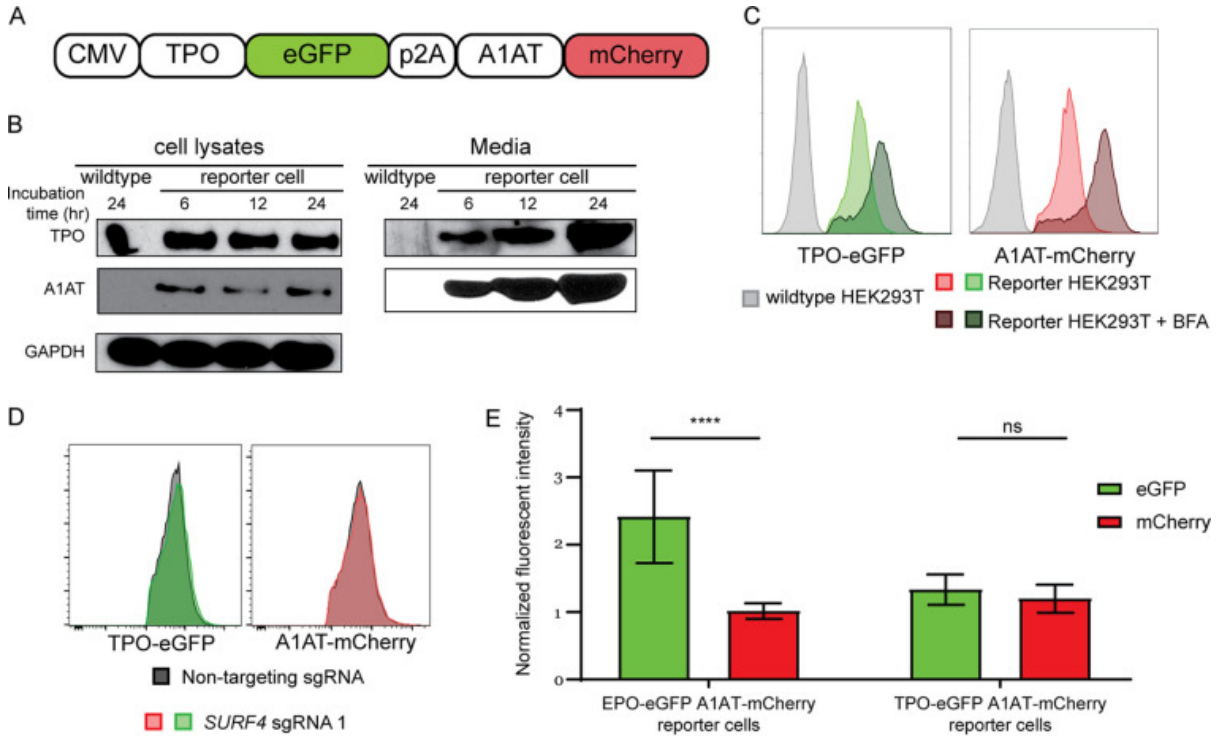

Figure 2–6 Thrombopoietin (TPO) secretion does not depend on SURF4. (A) A construct that expresses TPO-eGFP and A1AT-mCherry from the same CMV promoter was assembled and used to generate a reporter cell line stably expressing these two fusion proteins. (B) Intracellular and extracellular TPO-eGFP and A1AT-mCherry protein abundance was determined by Western blotting using anti-eGFP and anti-mCherry antibodies, respectively. (C) Inhibiting ER-to-Golgi transport with brefeldin A (BFA) leads to intracellular accumulation of TPO-eGFP and A1AT-mCherry, as measured by fluorescence intensity. (D) Flow cytometry histograms showing absence of intracellular accumulation of TPO following SURF4 deletion in HEK293T cells. (E) Quantification of cellular mean fluorescence intensity of TPO-eGFP and A1AT-mCherry in cells transduced with SURF4-targeting sgRNAs (n = 29). Results were normalized to mean fluorescence intensity of cells transduced with nontargeting sgRNAs. As a positive control, the same experiment was performed in parallel in reporter cell lines expressing EPO-eGFP and A1AT-mCherry (n = 48). ****, $P < 0.0001$; ns, not significant.

2.4.6 SURF4 promotes the secretion of endogenous EPO.

The experiments described above were performed in a heterologous cell line overexpressing EPO fused to either an eGFP or FLAG tag. To test the impact of *SURF4* deletion on the secretion of endogenous EPO, we transduced human HEP3B cells with *SURF4*-targeting sgRNAs or control sgRNAs. As a positive control, a sgRNA targeting *EPO* resulted in profound reduction of extracellular EPO level to almost an undetectable (0.45% of control) level (Fig. 2-7A). Disruption of *SURF4* in HEP3B cells using two independent sgRNAs resulted in a significant reduction (51.22% of control) of extracellular EPO levels compared to cells transduced with control sgRNAs (Fig. 2-7A), with no significant change in cellular *EPO* mRNA levels (Fig. 2-7B). Notably, *SURF4* mRNA levels increased by ~30% in cells treated with a prolyl hydroxylase inhibitor (Fig. 2-7C), suggesting that under conditions mimicking hypoxia, increased *SURF4* expression may contribute to enhanced EPO secretion.

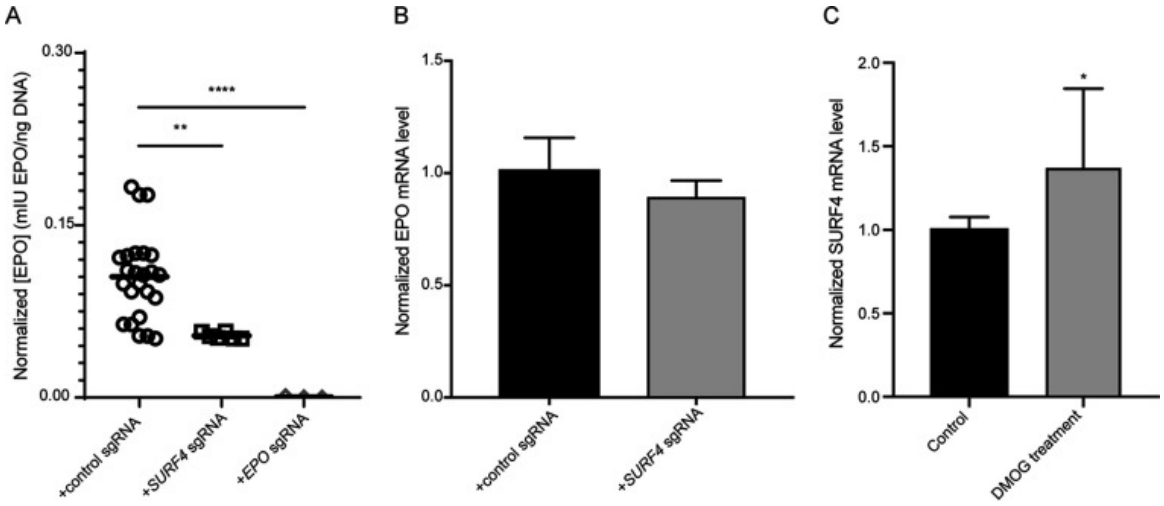

Figure 2–7 SURF4 deletion in HEP3B cells results in reduced extracellular secretion of EPO expressed from its endogenous genomic locus. mIU, milli-international units. (A) HEP3B cells were transduced with lentivirus expressing SURF4-targeting sgRNAs, control sgRNAs, or EPO-targeting sgRNA as a positive control. EPO expression from its endogenous regulatory elements was subsequently induced with CoCl2 and measured in the conditioned media by ELISA and normalized to the amount of cellular DNA (a surrogate of the total number of cells). (B) EPO mRNA expression by qRT-PCR in wild-type (n = 8) and SURF4-deleted (n = 8) HEP3B cells following CoCl2 treatment. (C) SURF4 mRNA expression by qRT-PCR in wild-type HEP3B cells with and without dimethyloxalylglycine (DMOG) treatment (n = 8 per condition). *. $P < 0.05$; **, $P < 0.01$; ****, $P < 0.0001$.

2.4.7 SURF4 overexpression promotes more efficient EPO secretion.

We next determined whether SURF4 overexpression promotes more efficient EPO secretion. We generated a lentivirus expressing equal amounts of SURF4 and Katushka2S (SURF4-p2A-Katushka2S) (Fig. 2-8A) and transduced it into HEK293 cells expressing EPO-FLAG from a tetracycline-inducible promoter. Cells with the highest (top 10%) and lowest (bottom 10%) SURF4 expression, as determined by Katushka2 fluorescence, were sorted by fluorescence-activated cell sorting (FACS). Following tetracycline administration, EPO level was found to be significantly increased in the conditioned media of cells overexpressing SURF4 compared to cells expressing low SURF4, with the reverse pattern observed intracellularly (Fig. 2-8B to E).

To assess the impact of SURF4 overexpression on the secretion of EPO expressed from its endogenous genomic locus, we performed the same experiment described above in HEP3B cells. EPO level was increased in the conditioned media of cells expressing high compared to low SURF4 levels (Fig. 2-8F), with no significant change in cellular *EPO* mRNA levels (Fig. 2-8G). Taken together, these results demonstrate that SURF4 overexpression promotes more efficient EPO secretion.

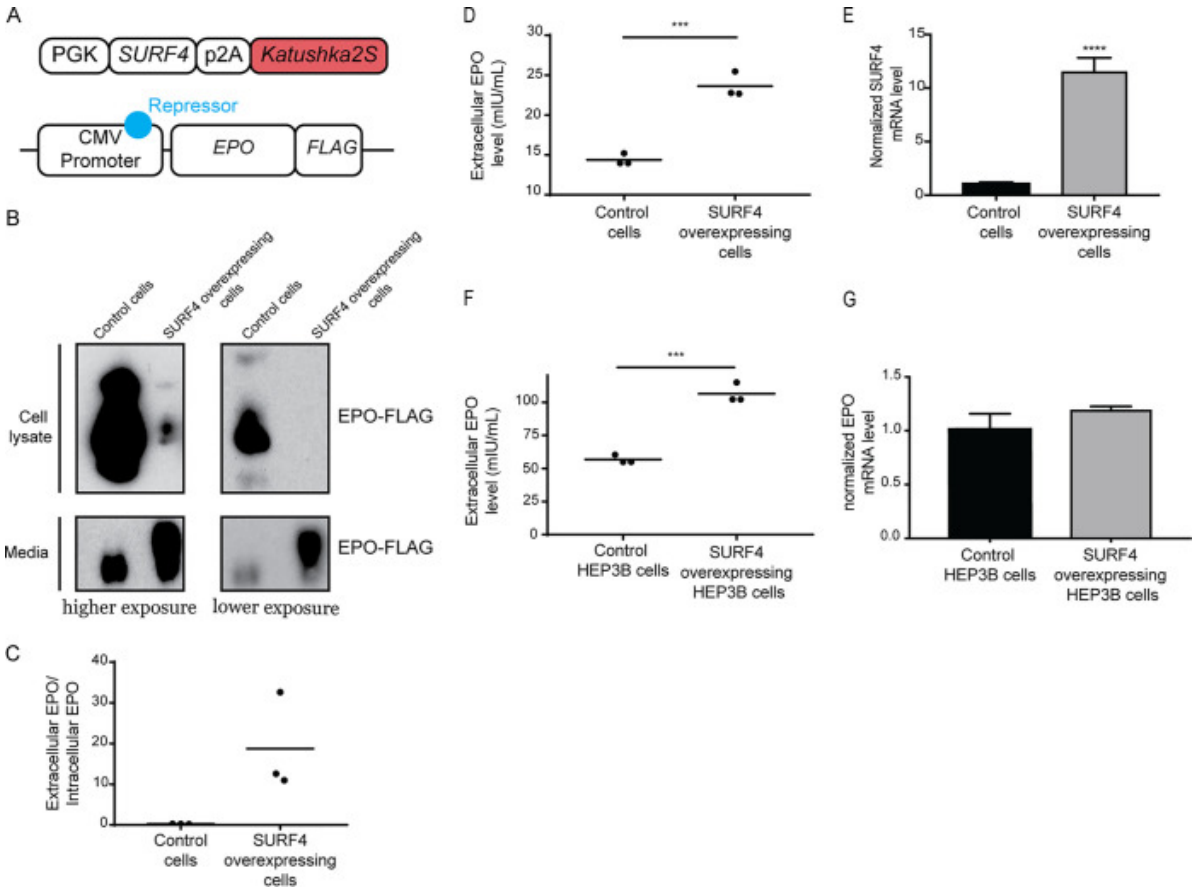

Figure 2–8 SURF4 overexpression leads to enhanced EPO secretion. (A) A lentiviral construct that expresses equal amounts of SURF4 and Katushka2S from the same phosphoglycerate kinase (PGK) promoter was assembled and transduced into HEK293 cells expressing EPO-FLAG from a tetracycline-inducible promoter. Cells with top 10% and bottom 10% Katushka2S fluorescence were FACS sorted, corresponding to cells overexpressing SURF4 and control cells, respectively. (B and C) Intracellular and extracellular EPO abundance following a 12-h tetracycline incubation was analyzed by Western blotting (using anti-FLAG antibody) (B), and quantification of densitometry of the ratio of extracellular/intracellular EPO was determined in three independent experiments (C). (D) The extracellular EPO level was also measured by ELISA. (E) SURF4 mRNA expression by qRT-PCR in wild-type ($n = 16$) and top 10% SURF4-expressing reporter cells ($n = 4$). (F) HEP3B cells overexpressing SURF4 (and control cells) were generated as described above. Following incubation with CoCl2, the extracellular EPO level was measured by ELISA. (G) EPO mRNA expression by qRT-PCR in wild-type ($n = 8$) and SURF4-overexpressing ($n = 3$) HEP3B cells following CoCl2 treatment. ***, $P < 0.001$; ****, $P < 0.0001$ by unpaired t test.

2.5 Discussion

In this report, we developed an unbiased genome-scale loss-of-function screen and identified *SURF4* as a regulator of intracellular EPO levels. Deletion of *SURF4* resulted in intracellular accumulation and extracellular depletion of EPO. Overexpression of SURF4 resulted in the reversed pattern. Consistent with the reported localization of SURF4 at the ER membrane [106, 118, 124], we found that disruption of *SURF4* resulted in accumulation of EPO in the ER and that EPO and SURF4 physically interact. Taken together, these results strongly suggest that SURF4 is the ER cargo receptor that mediates the efficient secretion of EPO.

The screen described above was performed in a cell line with heterologous overexpression of EPO under the control of a CMV promoter. Therefore, it was important to examine whether SURF4 facilitates the secretion of EPO when expressed at a more physiological level. Accordingly, we deleted *SURF4* in HEP3B cells that were induced to express EPO from its endogenous genomic locus and found that SURF4 also promotes EPO secretion under these conditions.

SURF4 is the mammalian ortholog of yeast Erv29p. Erv29p facilitates packaging of pro-alpha-factor in COPII vesicles promoting their ER-to-Golgi transport [117, 125, 126]. Erv29p recycles back from the Golgi apparatus to ER via recognition of its well- conserved dilysine sorting signal by the COPI coat [127]. In mammalian cells, only a handful of cargos (APOB, PCSK9, DSSP, AMLEX, and GH) have been shown to depend on SURF4 for efficient secretion [106, 118, 128]. However, a recent report demonstrated that mice with germ line deletion of *Surf4* exhibit early embryonic lethality [108] similar to *Caenorhabditis elegans* [118], suggesting the presence of

one or more SURF4- dependent cargos with a critical function during embryogenesis. Future studies aimed at identifying the repertoire of cargos that depend on SURF4 for secretion are essential. Equally important are future studies that examine the transcriptional landscape of SURF4-deficient EPO-producing cells (such as HEP3B cells).

A recently published report suggested that the cargo proteins that depend on SURF4 for secretion contain an N-terminal tripeptide "ER-ESCAPE" motif [128]. This motif is exposed following removal of the leader sequences and is recognized by SURF4 [128]. However, an N-terminal "ER-ESCAPE" motif with high SURF4 binding affinity is not present in EPO. Additionally, we found that EPO, but not TPO, depends on SURF4 for efficient secretion. However, TPO has an N-terminal motif with a better predicted SURF4 binding affinity than EPO. These results suggest that the N-terminal "ER-ESCAPE" motif may not be a generalizable determinant of SURF4 interaction for all SURF4-dependent cargos.

Soluble cargos are exported from the ER via the passive "bulk flow" or the concentrative "cargo capture" processes, with several lines of evidence supporting one route versus the other [90]. Though "bulk flow" and "cargo capture" are not mutually exclusive, this report provides support for the "cargo capture" model of EPO secretion. However, it is important to note that in our experimental conditions, ~50 to 70% of extracellular EPO is reduced in the setting of SURF4 deficiency. Therefore, the secretion of the remaining EPO depends on either bulk flow or one or more separate and unidentified receptors.

Recent developments in genome engineering using CRISPR/Cas9 technology have dramatically enhanced the potential and efficacy of comprehensive, high-throughput genetic screens [107, 111, 129-131]. Such strategies can be applied *in vitro* and *in vivo* to discover novel biologic insights. Our screen was designed to focus on posttranscriptional regulators of EPO by placing its expression under the control of a CMV promoter. Screening strategies similar to the one employed in this article and in a recently published report [106] might help identify additional ER cargo receptors for other soluble secreted proteins and shed more light into the extent of the contribution of "cargo capture" to recruitment of cargos into COPII vesicles.

Findings in this report may have important implications for erythropoiesis. EPO, the master regulator of erythropoiesis, is produced by specialized peritubular fibroblasts in the kidney. The transcriptional control of EPO via the hypoxia-inducible factor pathway has been well studied [75, 81, 132-135], culminating in the development of prolyl hydroxylase inhibitors, a class of compounds that increase EPO production at the transcriptional level via activation of the hypoxia-inducible factor [136-140]. These drugs are currently in clinical development, with several compounds in advanced phase 2 or 3 trials [141-145]; however, there are numerous potential concerns and adverse effects of these drugs, including possible increased risks of malignancy and autoimmune disease [146-148]. Similar to the transcriptional control of EPO, the intracellular signal transduction pathway downstream of the EPO receptor has also been well studied [82, 149, 150]. In contrast, much less is known about the molecular basis of EPO trafficking. Our findings suggest that modulating SURF4 may be effective for the treatment of disorders of erythropoiesis that are driven by aberrant EPO levels [86, 151-153]. Though a handful of other cargos depend on SURF4 for their secretion [106, 118, 128], with additional

cargos likely remaining to be identified, targeting *SURF4* exclusively in the EPO-producing cells might alter plasma EPO levels and therefore regulate erythropoiesis without affecting other SURF4-dependent cargos that are expressed in other cells. Alternatively, an inhibitor that specifically disrupts the SURF4-EPO interaction would also be expected to have no effects on other cargos that bind SURF4.

Recombinant human EPO (rhEPO) is used clinically for the treatment of anemia due to chronic kidney disease, chemotherapy, or ziduvidine. rhEPO is also used to reduce the requirement of allogeneic red blood cell transfusion following certain elective surgeries. Though the use of rhEPO is indicated in only a subset of the above clinical scenarios, the rhEPO market size was valued at ~7.4 billion U.S. dollars in 2016 [154]. In this report, we demonstrate that SURF4 overexpression results in enhanced EPO secretion. This approach could be applied to increase the efficiency of rhEPO production, which might translate into reduced costs of this drug.

Chapter 3 Loss-of-function CRISPR Screens Identify Novel Genes that Regulate Erythropoietin Production

3.1 Abstract

Erythropoietin (EPO) is a plasma glycoprotein that binds erythroid progenitors, stimulating their proliferation and differentiation. In adult humans, EPO is secreted into the circulation by hepatocytes and specialized kidney peritubular fibroblasts, contributing approximately 20% and 80% to the total plasma EPO level, respectively. EPO production is induced during hypoxia due to increased stability of the transcription factor hypoxia-inducible factor (HIF). However, HIF induction also results in increased expression of genes that promote angiogenesis and oncogenesis, which raises concerns about clinical strategies that promote EPO production via increased HIF levels. Here, we generated a reporter HEP3B cell line with homozygous insertion of coding sequences for p2A and eGFP at the endogenous *EPO* locus. This HEP3B EPO-p2A-eGFP cell line expresses equivalent levels of EPO and eGFP proteins, both translated from the same mRNA molecule. Using this cell line, we performed a genome-scale CRISPR knock-out screen with and without HIF activation to identify genes that regulate EPO production independent of HIF. This screen was followed by a secondary validation screen targeting the top 1,200 candidate regulators of EPO production nominated by the genome-scale screen. These studies identified ZNF574 as a novel HIF-independent regulator for EPO production. Under normoxia, deletion of *ZNF574* resulted in increased *EPO* mRNA and protein levels. Under hypoxic conditions, deletion of *ZNF574* also led to increased EPO production, more so than

wildtype cells exposed to hypoxia. RNA sequencing analysis showed that *ZNF574* deletion did not significantly impact the expression of most HIF-regulated genes. Taken together, these results suggest that ZNF574 regulates EPO production without activating the HIF pathway. These findings have important implications for several anemia disorders, particularly for anemia of chronic kidney disease, which results from impaired EPO production.

3.2 Introduction

EPO is the master regulator of erythropoiesis, the process by which red blood cells (RBCs) are produced. EPO binds its receptor expressed on erythroid progenitors in the bone marrow, promoting their survival, proliferation, and differentiation, resulting in increased RBC production [1-8].

During fetal development, EPO is produced predominantly by the liver [41-47]. Postnatally, the kidney becomes the major organ that produces ~80% of the circulating EPO, with a smaller contribution from the liver [28-30] [35-38]. Therefore, it is no surprise that patients with chronic kidney disease (CKD) exhibit anemia due to reduced EPO production.

While recombinant human EPO is clinically used to ameliorate the anemia observed in patients with CKD, it has its limitations, including elevated risk of cardiovascular events, the potential of developing EPO resistance and the risk of pure red cell aplasia [155]. EPO can elevate blood pressure and increase the risk of thrombosis, leading to potential heart attacks or stroke in susceptible patients [156, 157]. CKD patients are already at higher cardiovascular risk, and this added potential side effect from EPO requires careful consideration [158, 159]. Moreover, in

some patients, EPO resistance may develop, often due to inflammation or iron deficiency, reducing its efficacy [160]. Lastly, there may be a risk of pure red cell aplasia, a rare condition in which the body stops producing red cells in response to EPO therapy, usually associated with the production of anti-EPO antibodies [161].

Recently, prolyl hydroxylase inhibitors (PHD inhibitors), have been developed for the treatment of CKD anemia [162]. PHD inhibitors result in reduced degradation of the hypoxia-inducible factor (HIF) and therefore increased EPO production mediated by the binding of HIF to hypoxia-responsive elements (HRE) at the *EPO* locus [65-67]. Since PHD inhibitors result in increased intracellular levels of HIF, which also regulates cell proliferation and angiogenesis, the long-term use of PHD inhibitors is associated with theoretical concerns, which will only be addressed with long-term follow up of patients treated with these agents [163-167]. Therefore, increasing our understanding of EPO regulation, particularly its regulation independent of HIF, remains an unmet need in the field, with possible implications for the ~37 million US adults with CKD.

In this study, we aimed to identify novel regulators of EPO production by performing an unbiased genome-scale CRISPR knock-out screen, using a clonal HEP3B cell line that expressed eGFP from the endogenous *EPO* locus. To identify genes that regulate EPO production independent of HIF, this screen was performed with and without treatment with a PHD inhibitor. To validate the screen results, we generated a custom secondary library targeting the most promising 1,200 genes identified in the genome-scale screen. Validation experiments identified ZNF574 as a novel regulator of EPO production. Deletion of *ZNF574* results in increased *EPO* mRNA and protein levels, which appears to be mediated independent of HIF, since most HIF-

regulated genes were not impacted by *ZNF574* deletion. These findings have important implications for identifying novel strategies to increase endogenous EPO production without impacting HIF-regulated genes, therefore mitigating the risk of malignancies and autoimmune disease that result from sustained HIF activation.

3.3 Method

3.3.1 Cell culture

HEP3B and HEK293T cells were purchased from ATCC. HEP3B cells were cultured in MEM-alpha (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (Peak), 1% Glutamax (Gibco), and 1% penicillin-streptomycin (Gibco). HEK293T cells were cultured in DMEM (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (Peak) and 1% penicillin-streptomycin (Gibco). All cells were grown in a humidified 37 ^{o}C incubator with 5% CO2.

3.3.2 Generation of the HEP3B EPO-p2A-eGFP (HEG) reporter cell line

The DNA donor template used to knock-in p2A-eGFP at the *EPO* locus (prior to the stop codon of *EPO*) was assembled using the NEBuilder HiFi DNA assembly cloning kit (NEB). This construct contains the sequence encoding p2A-eGFP and 800 base pair homology sequence on either side of the insertion. sgRNA (AGGACAGGGGACAGATGACC) targeting the *EPO* sequence near the knock-in site was cloned into pSpCas9(BB)-2A-Purov2.0 (Addgene 62988), as described in chapter 2. Using Fugene HD transfection reagent (Promega), HEP3B cells were transfected with DNA donor template and sgRNA at 10:1 ratio. Transfected cells were selected with 1.5μg/mL puromycin (Sigma) for 3 days. Five weeks following puromycin selection, single

cells were sorted into 96-well plates pre-coated with bovine collagen (STEMCELL technology), using a MA900 Flow Cytometer (Sony). Single cell clones were expanded and analyzed for correct insertion of p2A-eGFP, using sanger sequencing.

3.3.3 Expansion and lentiviral preparation of the pLentiCRISPRv2 library

The pLentiCRISPRv2 whole genome CRISPR library was obtained from Addgene (Addgene #1000000048, a gift from Feng Zhang). The library was expanded by 16 electroporations (8 for each half library) into Endura electrocompetent cells (Lucigen), and plated on sixteen 24.5 cm bioassay plates (ThermoFisher Scientific). After 12-14 hour incubation at 37°C, colonies were harvested from agar plates, and pooled plasmids for each half library were isolated separately by Maxipreps using an EndoFree Plasmid Maxi Kit (Qiagen). Viral particles were obtained by co-transfection (using Lipofectamine LTX with PLUS reagent (ThermoFisher Scientific) of 17µg of PAX2 (addgene #12260, a gift from Didier Trono), 11.3 µg of pCMV-VSV-G (Addgene #8454, a gift from Robert Weinberg), and 11.3 µg of each half library into HEK293T cells plated at 80~90% confluency in six T225 tissue culture flasks (ThermoFisher Scientific). Twenty-four hours post-transfection, media was changed, and 12, 24, 36 hours post media change, the viral supernatant was harvested. Cell debris were removed from viral supernatant by 500g centrifugation for 5 minutes. The viral supernatant was then pooled, aliquoted, snap-frozen by liquid nitrogen and stored at -80 ^{o}C.

3.3.4 Design and synthesis of custom CRISPR library

Candidate genes from the genome-scale screen were sorted by their relative ranking using MAGeCK analysis. From each screen condition (with and without DMOG), 349 positive and negative candidate regulators were selected. Using GPP sgRNA designer (Broad institute), 10 optimized sgRNAs were generated per gene when possible, and 118 non-targeting sgRNA controls were included. Flanking sequences were appended to sgRNA sequences as priming site for cloning. Synthesized pooled oligonucleotide were purchased from Twist Bioscience and was cloned into Lenti-CRISPRv2 as described in chapter 2. The custom library was expanded by 3 electroporations into Endura electrocompetent cells (Lucigen). Maxipreps, lentiviral packaging, and lentiviral harvesting were done in the similar fashion as the whole-genome library, mentioned above. Library diversity was assessed by deep sequencing using Illumina Miseq, confirming adequate expansion of all 12000 sgRNA and a $90^{th}/10^{th}$ percentage read ratio of 2.99.

3.3.5 CRISPR/Cas9 Genome-scale loss-of-function screens

For each replica of the genome-scale screen, a minimum of 132 million reporter cells were plated in 15-cm tissue culture dishes (Corning) at 4.16 million cells per dish. On day 0, cells were transduced with the lentiviral library (with 8µg/mL polybrene, Sigma) at a multiplicity of infection (MOI) of ~0.3. Twenty-four hours post-viral transduction, transduced cells were selected with puromycin (1µg/mL, sigma) for 4 days. Between day 4 and day 12, cells were kept at a logarithmic phase of growth and passaged every 2-3 days, while ensuring more than 37 million cells were in culture at all times, to maintain ~300x coverage. Thirteen-days post-transduction, cells were split into two groups. One group was left untreated. The other group was treated with 200mM DMOG for 24 hours. On day 14, ~80 million cells from each group were isolated from tissue culture dishes using 0.25% trypsin (Gibco), pelleted by centrifugation (350g,

4oC, 5 min), resuspended in cold PBS+2% FBS, and filtered through a 35 μm mesh into flow cytometry tubes (corning). Cells were maintained on ice until sorting. Cells with top or bottom 10% eGFP fluorescence (~3 million cells/population; ~24x coverage) were sorted using MA900 Flow Cytometer (Sony) and collected into 15 mL polypropylene tubes (Cellstar) containing media. Genomic DNA was extracted using a DNeasy Blood & Tissue kit (Qiagen), and integrated lentiviral sgRNA sequences were amplified by a two-step PCR reaction (20 cycles and 14 cycles, respectively) as previously described [168]. DNA amplified by each PCR reaction was purified using Zymo-Spin V column (Zymo research), DNA-binding buffer (Zymo), DNA wash buffer (Zymo) and Ultra-pure distilled water (Invitrogen). The DNA integrity and yield was analyzed by Bioanalyzer (Agilent) and qPCR (QuantStudio), respectively. sgRNA amplicons were sequenced using NextSeq 75 cycles HO (Illumina) with samples pooled. On average, 23.5 million reads were generated for each sorted cell population of each screen. Overall, 98% of the reads had a per sequence quality score (phred-based base quality score) of greater than 30. 104,331 sgRNA sequences were mapped and identified (along with the barcode corresponding to each cell population of each replicate) using a custom Perl Script as previously described . Enrichment at the sgRNA and gene levels was analyzed using MAGeCK.

3.3.6 CRISPR/Cas9 custom loss-of-function screens

For the custom loss-of-function screen, a minimum of 132 million reporter cells were plated and the work flow was similar to the one described for the genome-scale screen described above. On the day of sorting, on average, 2.35 million cells with either top or bottom 10% eGFP fluorescence were collected. DNA extraction, sgRNA sequence amplification, PCR purification and Illumina sequencing were done as described above for genome-scale screen.

3.3.7 Flow Cytometry analysis

HEP3B cells were detached from tissue culture dish with 0.25% trypsin (Gibco), washed with PBS, collected by centrifugation (350g, 5 min, 4°C) and resuspended in cold PBS with 2% BSA and 10mM HEPES (Invitrogen). HEP3B cells were analyzed by ZE5 cell analyzer (BioRad). Mean fluorescent intensity was calculated by ZE5 cell analyzer.

3.3.8 Deletion of candidate genes with CRISPR/Cas9

sgRNAs targeting candidate genes and non-targeting sgRNAs were cloned into the pLentiCRISPRv2 plasmid as previously described [168]. Lentiviral infection, selection of the transduced cells, and the treatment with DMOG were performed as described for the CRISPR screen above. eGFP fluorescence intensity was measured by flow-cytometry, as described above.

3.3.9 Insertion deletion (indel) efficiency

Genomic DNA of HEP3B cells was isolated by QuickExtract DNA Extraction Solution (Biosearch Technologies), following manufacturer's protocol. PCR was performed using primers aimed to amplify the genomic DNA sequence that contains the expected sgRNA cut-site. For *ZNF574*, this was done using primers 1: AGAGGAATGGGGTGAGGTGA; primer 2: CGCGACACTTGTAGGGGTAT. The PCR product was sequenced by Sanger sequencing, and chromatograms were analyzed by ICE CRISPR Analysis Tool (Synthego) to calculate the indel efficiency.

3.3.10 EPO ELISA

On day 0, 30,000 cells were seeded per well in 96-well plates. One day after seeding, cells were treated with DMOG or left untreated for 24 hours and conditioned media was collected and centrifuged (500g, 5 min, 4 ^{o}C) to remove potential cell debris. EPO level was measured using LEGEND MAX Human Erythropoietin ELISA kit (Biolegend), following the manufacturer's protocol. The extracellular EPO level was normalized to the total cell number, measured by MTT cell proliferation assay (Sigma).

3.3.11 Quantitative RT-PCR (qRT-PCR)

RNA was isolated using RNeasy kit (Qiagen). cDNA was generated using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems), following the manufacturer's protocol. qRT-PCR was performed using Power SYBR Green PCR master mix, and QuantStudio 7 real-time PCR system (ThermoFisher). Samples were analyzed in triplicates and beta-actin was used as internal control. The fold change in *EPO* mRNA level was measured by ΔΔCt method. Primers used were as follows: *EPO* Forward primer: CGGAGATGGGGGTGCAGA, *EPO* Reverse primer: TGGGAGGCCCAGAGGA; *ACTB* Forward primer: CACCATTGGCAATGAGCGGTTC, *ACTB* Reverse primer: AGGTCTTTGCGGATGTCCACGT.

3.3.12 Statistical analysis

CRISPR screen analysis was performed by MAGeCK. Differences in mean fluorescence intensity in eGFP levels, extracellular EPO levels, and *EPO* mRNA levels were compared by one-way ANOVA.

3.4 Results

3.4.1 Generation of a HEP3B cell line that reports the EPO production level.

To identify novel proteins that regulate EPO production, we generated a HEP3B reporter cell line with homozygous insertion of coding sequences for p2A and eGFP at the endogenous *EPO* locus (Fig. 3-1A). This HEP3B EPO-p2A-eGFP cell line, referred to as HEG, expresses equal amounts of EPO and eGFP proteins, translated from the same mRNA molecule. Therefore, proteins that regulate *EPO* transcription or translation are expected to equally affect the intracellular eGFP level. Since the eGFP protein expressed by the HEG cell line is not fused to EPO, proteins that regulate EPO secretion or degradation are not expected to affect the intracellular eGFP level.

In normoxia, the HEG cell line produces EPO but at a low basal level (Fig 3-1B). Following incubation of HEG cells with 200uM of the PHD inhibitor DMOG for 24 hours (selected based on treatment time course and dose response curve analyses), EPO production increased significantly compared to baseline, as determined by ELISA (Fig. 3-1D) and flow cytometry (Fig. 3-1E). Deletion of the hypoxia responsive element (HRE) located downstream of the *EPO* coding sequence (Fig. 3-1F) resulted in abrogation of the DMOG-mediated increase in EPO production (Fig. 3-1G). These findings validate the utility of the HEG cell line to study the regulation of *EPO* expression.

Fig. 1

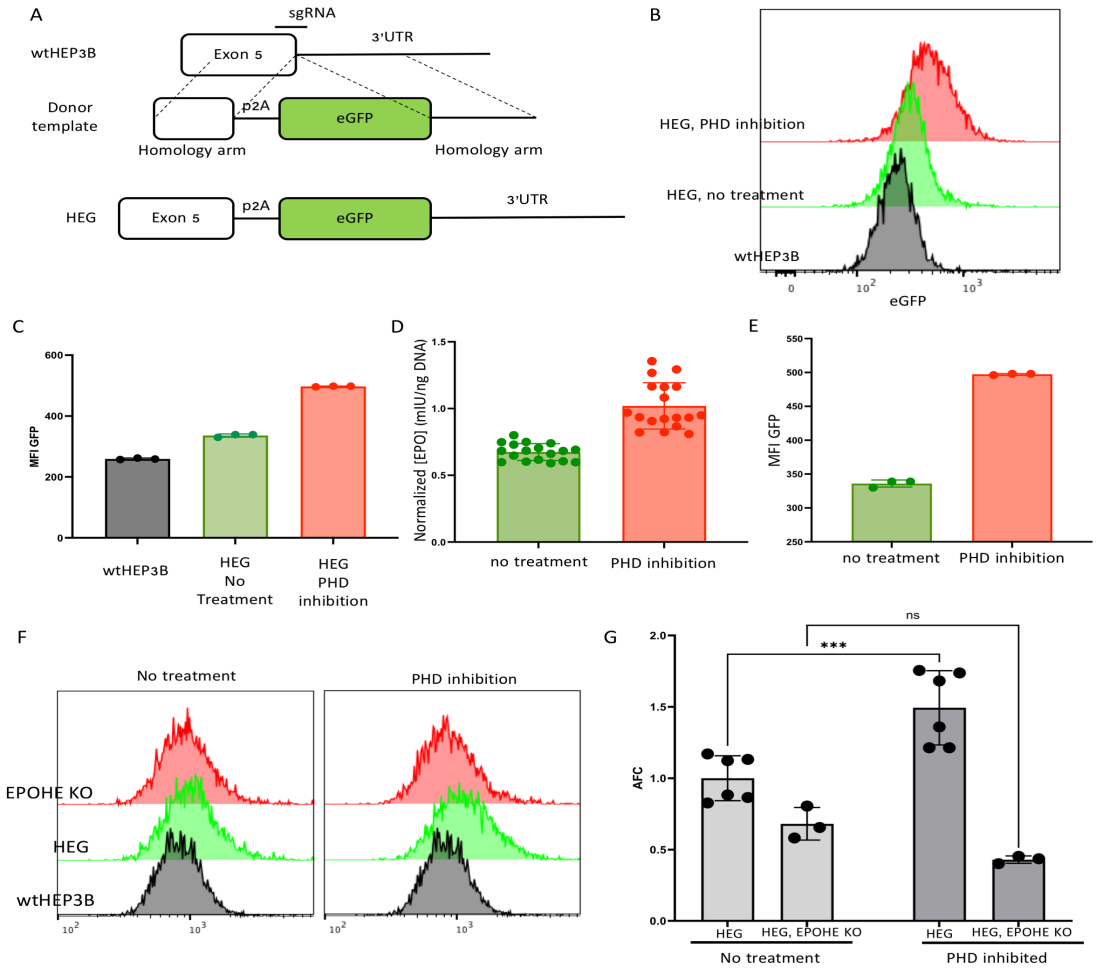

Figure 3–1 A reporter HEP3B cell line expressing EPO p2A GFP from the *EPO* locus. (A)Schematic representation of the strategy for insertion of p2A-eGFP at the genomic locus of *EPO*, right before the *EPO* stop codon. (B) Representative fluorescent histogram demonstrating eGFP accumulation in HEG cells following treatment with DMOG. (C) Quantification of intracellular mean fluorescence intensity (MFI) in three independent experiments. (D) EPO production by HEP3B cells after DMOG treatment, measured by ELISA (n=18). (E) Intracellular eGFP level in HEG after DMOG treatment, measured by flow cytometry (n=3). (F) Flow cytometry histograms comparing the eGFP signal in HEG with and without deletion of hepatic *EPO* enhancer. (G) Adjusted fold change (AFC) of MFI GFP from (F), demonstrating loss of hypoxia-mediated induction of *EPO* in HEG following deletion of hepatic *EPO* enhancer (n=3), normalized to HEG, no treatment.

3.4.2 A genome-scale CRISPR/Cas9 screen for regulators of EPO expression

To identify novel proteins that regulate EPO production, we developed a genome-scale CRISPR/Cas9 screen using the HEG reporter cell line, which provides a quantifiable and selectable readout of EPO production. We mutagenized ~37 million HEG cells with the lentivirally packaged human GeCKOv2 CRISPR knockout library, which delivers spCAS9, a pooled collection of 6 sgRNAs targeting nearly every gene in the genome, and a puromycin resistance cassette. Transduction was performed at a low multiplicity of infection (~0.3) so that most infected cells receive 1 sgRNA to delete 1 gene in the genome. Following selection of transduced cells with puromycin, cells were recovered for 8 days and were subsequently either treated with DMOG for 24 hours or left untreated (Fig. 3-2A). Cells with increased (top ~7%) or decreased (bottom ~7%) eGFP fluorescence were isolated, and integrated sgRNA sequences were quantified by deep sequencing (Fig. 3-2A). This screen was done in biological quadruplicates for the no treatment arm and in biological triplicates for the DMOG-treated arm.

To identify genes that impact *EPO* expression, sgRNAs were analyzed for their enrichment or depletion in the eGFP high compared to the eGFP low cell populations (Fig 3-2A). Since DMOG stimulates *EPO* expression via activation of the HIF pathway, genes for which sgRNAs are enriched/depleted in the eGFP high compared to the eGFP low DMOG-treated cells likely impact *EPO* expression independently of the HIF pathway. In contrast, genes for which sgRNAs are enriched/depleted in the eGFP high compared to eGFP low untreated cells may impact *EPO* expression in either a HIF dependent or independent manner.

Based on gene level analysis, sgRNAs targeting *HIF1A*, *EPAS1*, and *HIF3A* (which encode HIF1α, HIF2α, and HIF3α, respectively) were not enriched in the eGFP-low compared to the eGFP-high cell population (Fig 3-2B), likely due to overlapping roles for these genes in stimulating EPO production. Similarly, and likely for the same reason, sgRNAs targeting *EGLN2*, *EGLN1*, and *EGLN3* (which encode PHD1, PHD2, and PHD3, respectively) were not enriched in the eGFP-high cell pool (Fig 3-2B).

In contrast, as expected, sgRNAs targeting *VHL* were found to be significantly enriched in the eGFP-high population (Fig 3-2C). sgRNAs targeting genes known to promote *EPO* expression (*HNF4A*, *WT1*, *RXRA*, and *CREBBP*) were significantly enriched in the eGFP-low compared to the eGFP-high cell population under either DMOG-treated or untreated conditions. These findings provide confidence in the screen results.

Fig. 2

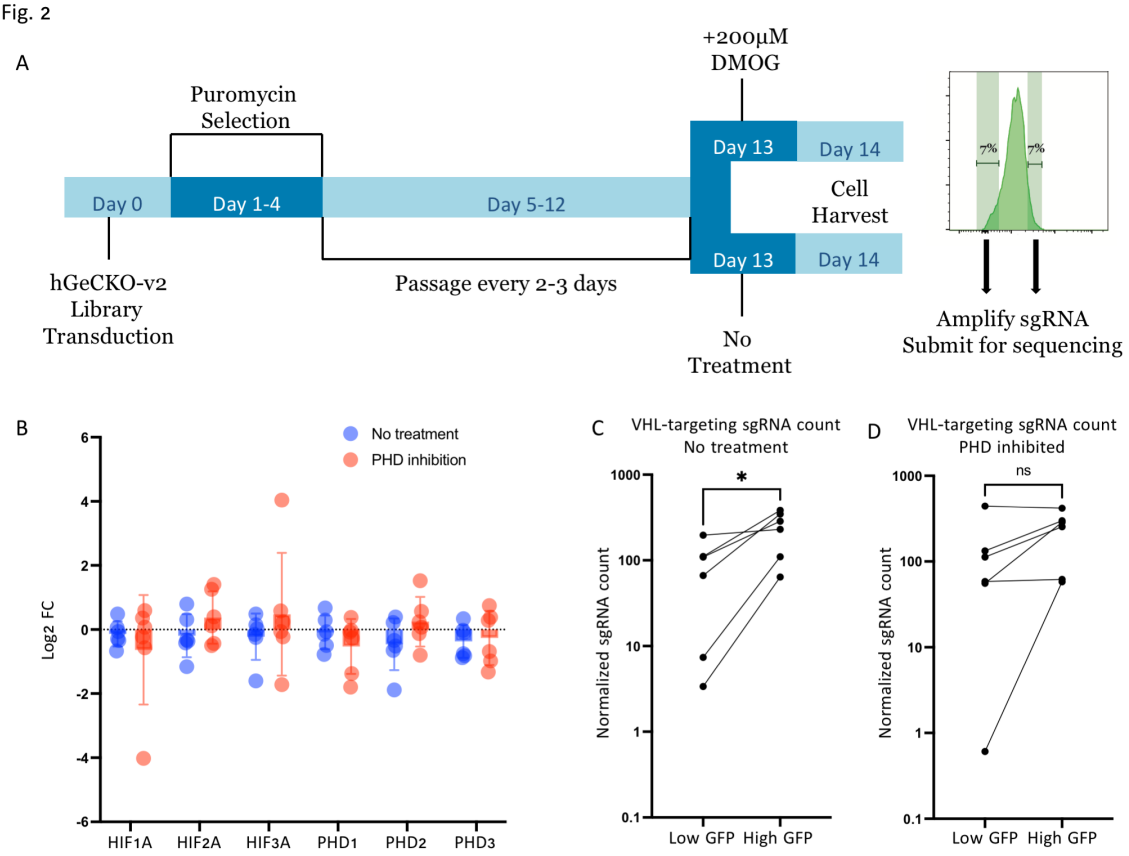

Figure 3–2 A genome-scale CRISPR/Cas9 screen for regulators of EPO expression. (A) Screen strategy. Twenty-four hours following transduction of the CRISPR library, puromycin selection was applied for 3 days. AT day 13, cells were either treated with DMOG or left untreated. Twenty-four hours later, cells with top or bottom 7% eGFP fluorescence were isolated. sgRNA abundance was then determined in each cell population. (B) Log 2 fold change in sgRNA abundance between eGFP high and eGFP low cell populations for the three HIF isoforms and three PHD isoforms, under no treatment condition or PHD inhibition. (C, D) Normalized counts for the six *VHL*-targeting sgRNAs including in the library in both eGFP high and eGFP low cell populations, for screen performed under (C) no treatment condition and (D) DMOG treated conditions.

3.4.3 Secondary screen for regulators of EPO production

To refine our genome-scale screen findings, we selected the top 349 genes for which sgRNAs were enriched in either of the 4 conditions: eGFP high compared to eGFP low and eGFP low compared to eGFP high cell populations, in either DMOG-treated or untreated conditions. Due to some overlap in the genes across conditions, the total number of genes included was 1,200 (instead of 1,396). We generated a custom sgRNA library that contains 10 sgRNAs targeting each of the 1,200 genes when possible and 118 non-targeting sgRNAs as control. The sgRNAs were cloned en masse into the Lenti-CRISPRv2 plasmid. Massive parallel sequencing of the plasmid pool confirmed presence of all sgRNA sequences (Fig 3-3A) with minimum library skewing (Fig 3-3B), as demonstrated by $90^{th}/10^{th}$ percent read count of 2.9. The secondary library was packaged into lentiviral particles and the secondary screen was performed as described above for the genome-scale screen (Fig 3-2A).

Owing to the considerably (~10 fold) smaller size of the secondary library compared to the genome-scale library, the secondary screen was performed with a substantial average coverage of ~250x (i.e. each sgRNA was represented ~250 times on average) throughout the entire screen. To achieve such coverage, ~43 million cells were maintained in culture, allowing 3 million cells with top 7% and 3 million cells with bottom 7% EPO production to be sorted, thus maintain the ~250x sgRNA coverage in each sorted cell population. The high coverage achieved with the secondary screen combined with the large number of sgRNAs/gene included in the secondary library were designed to improve our ability to distinguish true hits from background noise. Sequencing generated 25,603,343 reads in total, with an average of ~213 reads per sgRNA.

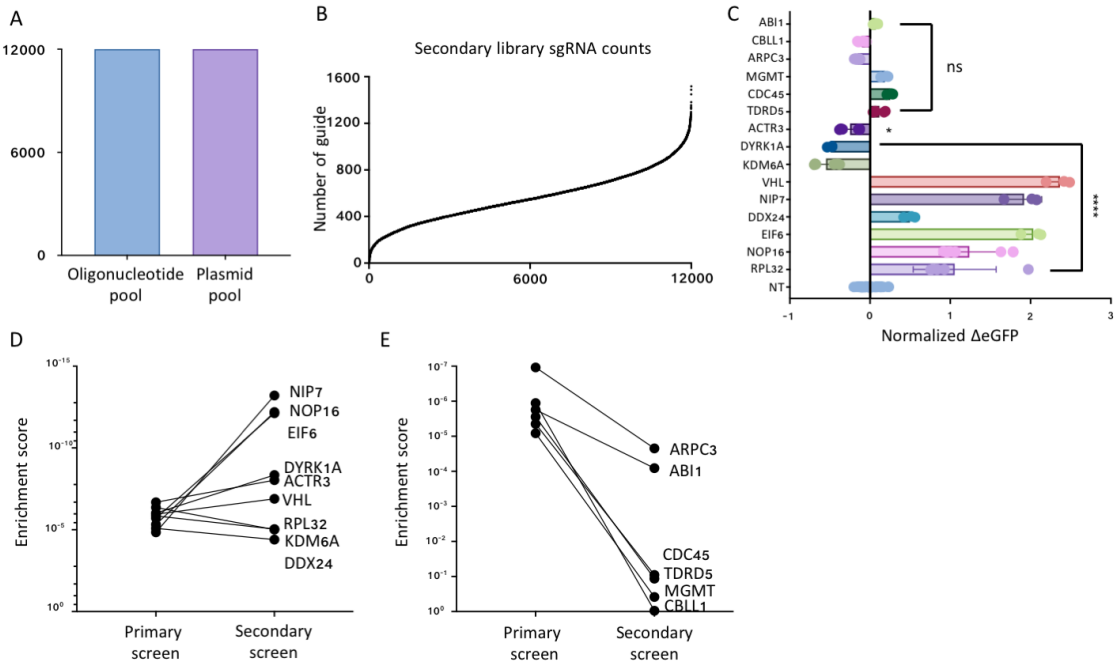

Figure 3–3 Secondary screen for regulators of EPO production. (A) All sgRNA guides that were synthesized were identified by sequencing. (B) Normalized sgRNA counts for all sgRNAs in the plasmid pool demonstrating no significant skewing in sgRNA expansion. (C) Changes in eGFP MFI for high-ranking candidate genes, using the most promising sgRNA identified in the screen. n = 3 for each candidate gene and n =12 for non-targeting sgRNAs. Error bar represent standard deviation. (D) Comparison of MAGeCK gene level enrichment scores for validated candidates in primary versus secondary CRISPR screens for EPO production. (E) Comparison of MAGeCK gene level enrichment scores for the candidates that failed to validate in primary versus secondary CRISPR screens.

3.4.4 Comparison between the genome-scale and secondary screens

To determine if the secondary screen fared better than the genome-scale screen at separating true hits from false positive hits, we chose the top 9 genes (excluding *SYS1*, due to its known function in GFP secretion) from the genome-scale library screen for which sgRNAs were enriched in the eGFP high compared to the eGFP low population and the top 6 genes for which sgRNAs were enriched in the eGFP low compared to the eGFP high population. Using an sgRNA targeting each of these 15 candidate genes, 6 of the 9 putative negative regulators and 3 of the 6 putative positive regulators of EPO production were validated (Fig 3-3C).

Gene level enrichment scores were determined by MAGeCK, reflecting the relative abundance of sgRNAs between eGFP high and low cells populations as well as the concordance across various sgRNAs targeting the same gene. Enrichment scores for the 9 genes that validated as regulators of EPO production were overall improved in the secondary screen compared to the genome-scale screen (Fig 3-3D). Similar analysis showed that MAGeCK gene level enrichment scores for the 6 genes that were proven to not impact EPO production were overall significantly reduced in the secondary screen compared to the genome-scale screen (Fig 3-3E). These results demonstrate that the secondary screen appears to better separate true positive from false positive hits compared to the genome scale CRISPR screen, as expected.

3.4.5 Analysis of the secondary CRISPR screen

Using a false discovery rate (FDR) cutoff of 1%, gene-level analysis identified 100 genes that serve as negative regulators of EPO production (i.e. whose targeting was associated with

increased EPO production, Fig 3-4A and 3-4B), under either untreated or DMOG-treated conditions. As demonstrated in Fig 3-4C, a high degree of concordance was observed for negative regulators of *EPO* expression between both screens, performed with or without DMOG treatment (Fig 3-4D).

Similarly, gene-level analysis identified 79 genes that positively regulate EPO production (i.e. whose targeting was associated with reduced EPO production, Fig 4A and 4B), under either DMOG treated or untreated conditions at FDR <1%. A high level of concordance was also observed for positive regulators of *EPO* expression between screens performed with and without DMOG (Fig 3-4E).

To identify transcription factors that regulate EPO production, we first selected the 4 most highly ranked genes in the screen that encode DNA-binding proteins, and deleted each candidate gene or *VHL* as positive control in the HEG cell line using CRISPR/Cas9. Compared to cells transduced with non-targeting sgRNAs, deletion of *VHL* resulted in increased EPO in normoxia, with no further increase in EPO production in *VHL* deleted cells following DMOG administration (Fig 3-4F). These results are expected, since *VHL* deletion and DMOG administration induce EPO production via the same mechanism, mediated by HIF.

In normoxia, deletion of *MLL2* resulted in reduced EPO production, while deletion of *TAF1C* or *ZNF574* resulted in increased EPO production (Fig 3-4G). In contrast, under DMOG treatment conditions, deletion of *MLL2*, *TAF1C*, *ZNF574*, or *RFX7* resulted in increased EPO production compared to control cells transduced with non-targeting sgRNA (Fig 3-4H).

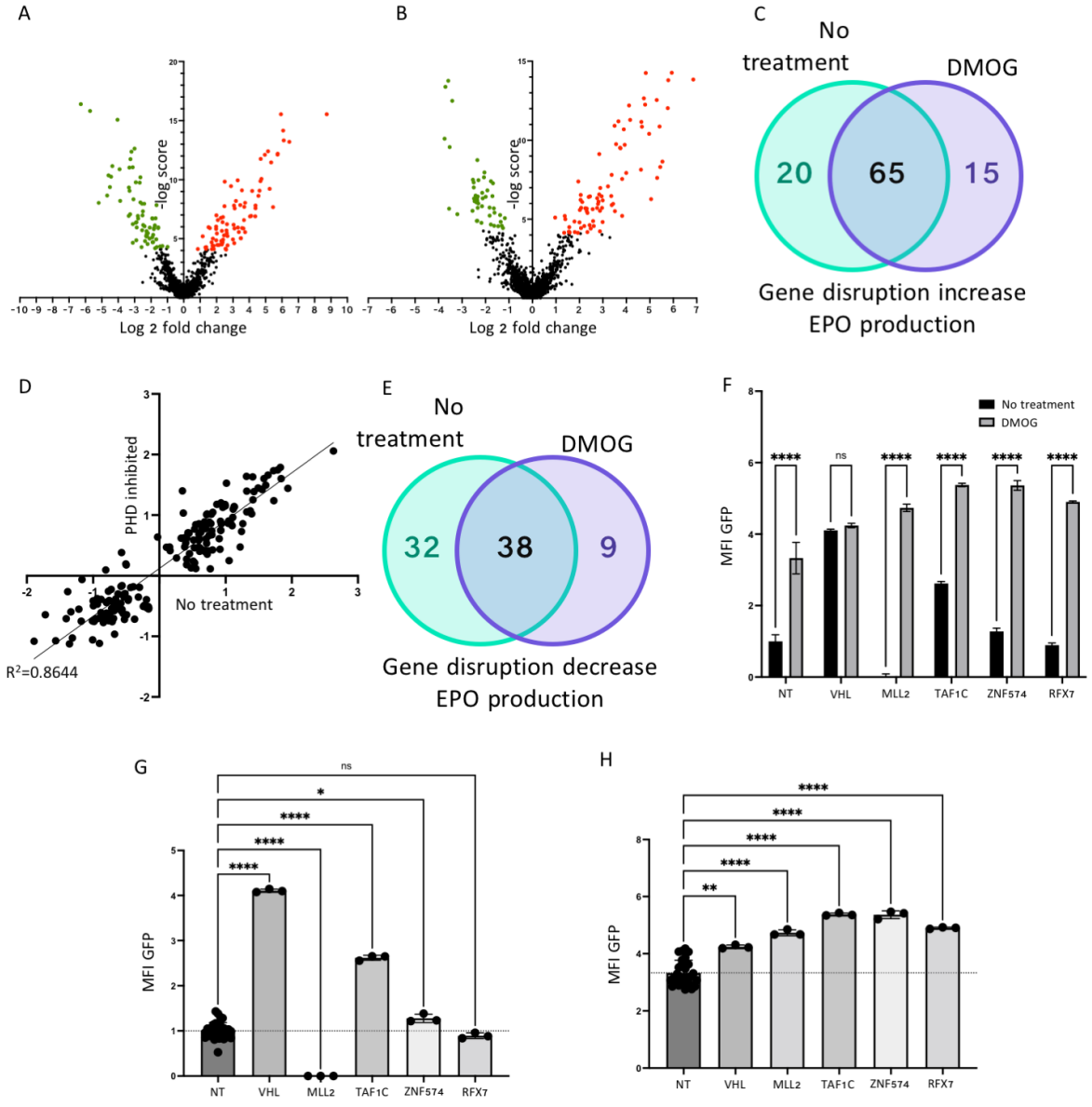

Figure 3–4 Analysis of the secondary CRISPR screen. (A,B) Volcano plots representing gene-level analysis, with the log2 fold changes of gRNAs abundance for each gene in the secondary library on the x-axis and the statistical significance on the y-axis. Volcano plots in no treatment condition is shown in panel (A) while data with DMOG treatment is shown in panel (B), with genes identified with FDR<1% displayed in green and red. (C, E) Venn diagrams of gene identified whose targeting was associated with increased (C) or reduced (E) EPO production under no treatment or DMOG treated conditions. (D) Scatter plot shows log2 fold enrichment for genes with FDR<1% in the no treatment (x-axis) and PHD inhibited (y-axis) conditions. The log2 fold enrichment (in eGFP high versus eGFP low cells) was calculated by MAGeCK analysis. A good correlation was observed between both conditions. The best fit line has R^2=0.8644. (F) Comparison of EPO induction in HEG cells (measured by mean GFP fluorescence) between untreated and DMOG treatment conditions following deletion of certain (indicated on the x-axis) versus control cells transduced with non-targeting (NT) sgRNA. (G,H) Intracellular mean eGFP flourescence intensity in HEG cells deleted for genes (indicated on the x-axis) versus control cells transduced with non-targeting (NT) sgRNAs, under no treatment (G) or DMOG (H) conditions.

3.4.6 ZNF574 deletion leads to increase in EPO production

ZNF574 deletion resulted in the highest increase in EPO production with HIF activation. Therefore we selected *ZNF574* for further evaluation. To rule out an off-target effect of the *ZNF574*-targeting sgRNA, we generated three additional *ZNF574*-targeting sgRNAs (sgRNAs 2-4). All *ZNF574* targeting sgRNAs were found to be highly efficient by TIDE analysis, resulting in insertions/deletions rates between 89% and 97%. *ZNF574* deletion using all 3 sgRNAs resulted in a ~2-fold increase in intracellular EPO levels ($p<0.0001$) at normoxia (Fig3-5A). Therefore, we have validated that *ZNF574* deficiency results in increased EPO production using 4 independent sgRNAs.

Notably, under hypoxic condition (2% oxygen), deletion of *ZNF574* resulted in a more profound increase in EPO level compared to i) cells receiving non-targeting sgRNAs under DMOG conditions and ii) cells deleted for *ZNF574* under normoxia conditions (Fig. 3-5B), suggesting that in contrast to VHL, ZNF574 regulates EPO production independent of HIF.

We next deleted *ZNF574* in HEP3B cells and measured the cellular *EPO* mRNA level by qRT-PCR as well as the EPO protein level in conditioned media by ELISA. *ZNF574* deletion resulted in a significant increase in *EPO* mRNA levels under hypoxia but not under normoxia (Fig 3-5C and 3-5D). In normoxia, the amount of EPO protein in conditioned media is below the detection limit of the ELISA kit. In hypoxia, EPO protein is higher in conditioned media of cells deleted for *ZNF574* compared to cells transduced with non-targeting sgRNAs (Fig 3-5E). Altogether,

these results demonstrate that *ZNF574* deletion results in increased EPO production at the transcriptional level.

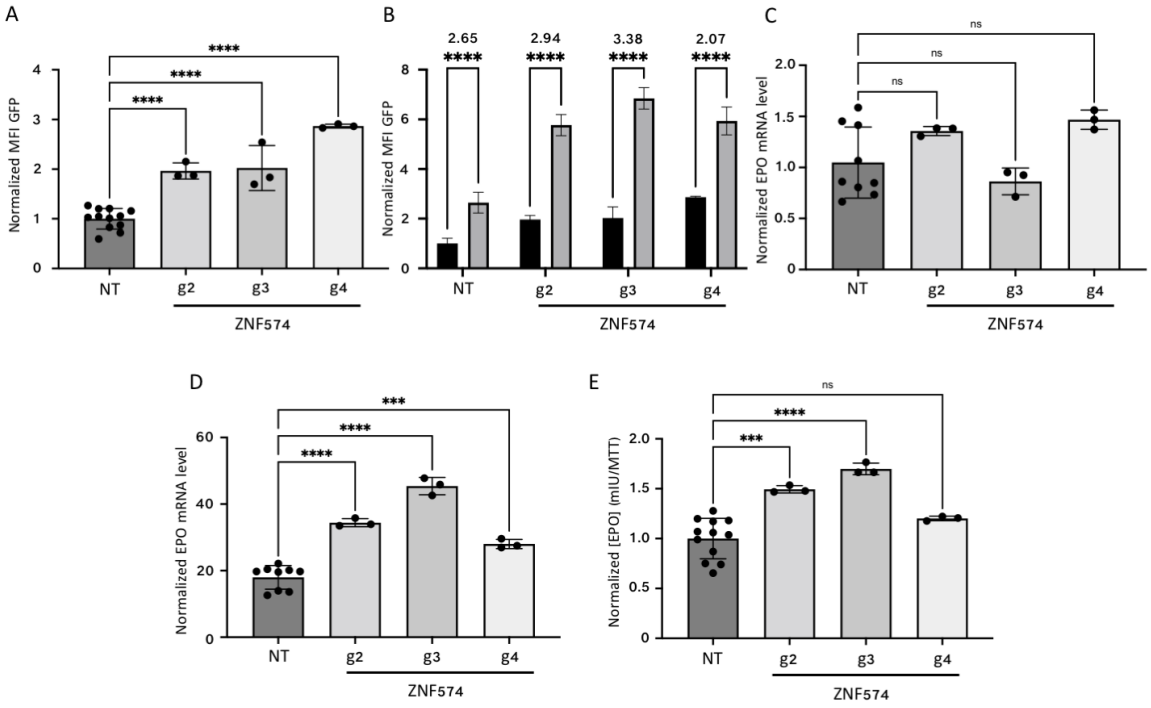

Figure 3–5 *ZNF574* deletion leads to increase in EPO production. (A) A comparison of mean eGFP fluorescence intensity in cells treated with control non-targeting (NT) sgRNA versus cells treated with three distinct *ZNF574*-targeted sgRNAs in normoxia. (B) Comparison of EPO induction in HEG cells (measured by mean GFP fluorescence) between normoxia (black bars) and hypoxia (gray bars) conditions for cells transduced with ZNF574-targeting sgRNAs or non-targeting sgRNA. (C, D) *EPO* mRNA level measured by qRT-PCR in cells transduced with non-targeting control sgRNA or three indepedent *ZNF574*-targeting sgRNA, under normoxia (C) and hypoxia (D). (E) Extracellular EPO level in conditioned media from cells transduced non-targeting sgRNA or *ZNF574* targeting sgRNA; cells were incubated in hypoxia for 24-hour prior to media collection.

3.4.7 Genes regulated by ZNF574

To determine the repertoire of genes that are regulated by *ZNF574*, we compared the transcriptomes of *ZNF574* null HUDEP2 cells to WT control cells, both under normoxia and hypoxia (2% O2) conditions.

Under normoxic conditions, using an average fold change cutoff of 2-fold and adjusted p-value <0.05, we found that the expression of 80 genes is increased while the expression of 45 genes is decreased in *ZNF574* null cells (Fig 3-6A). Using similar statistical cutoffs, we found that under hypoxia, the expression of 235 genes is increased, while the expression of 94 genes is decreased in *ZNF574* null cells (Fig 3-6B). A high concordance is observed between genes regulated by ZNF574 under conditions of normoxia and hypoxia (Fig3- 6C). As previously shown, *EPO* mRNA level is increased more significantly in *ZNF574* deficient cells exposed to hypoxia (Fig 3-6D).

We next evaluated the expression of *HIFα* and known HIF-regulated genes in *ZNF574* null versus WT cells. The expression of *HIF1A*, *EPAS1*, and *HIF3A* (which encode HIF1α, HIF2α, and HIF3α, respectively) was not significantly different between *ZNF574* deleted cells and WT controls. Additionally, while the expression of a minority of HIF-regulated genes was mildly increased in ZNF574 deficient cells, the expression of the majority of HIF-regulated genes, such as *LDHA*, *IGF2*, and *CA9*, was unchanged (or even mildly reduced) in ZNF574 deficient versus WT cells (Fig 3-6E). Taken together, these data suggest that ZNF574 deficiency results in increased EPO without affecting all HIF regulated genes, suggesting that this is mediated independent of the HIF pathway, though this remains to be definitively proven.

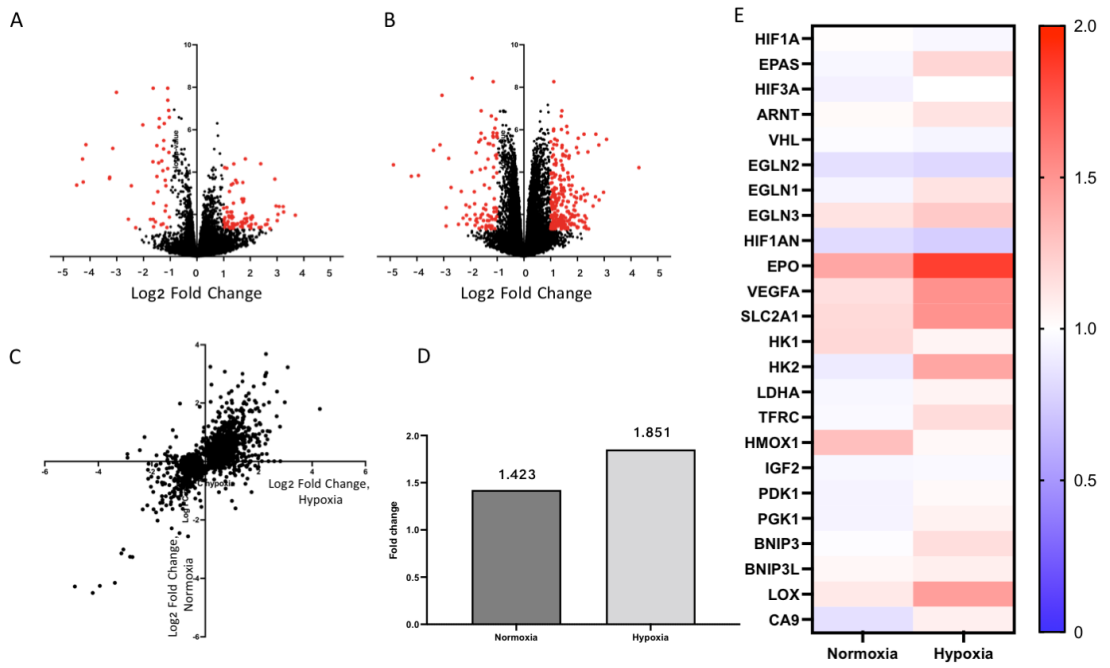

Figure 3–6 Genes regulated by ZNF574. Volcano plots comparing gene expression in *ZNF574* null versus WT HEP3B cells under (A) normoxia and (B) hypoxia. Each dot in plots represents an individual gene, genes that are significantly regulated are highlighted in red, marked by an absolute fold change of 2 or higher and a adjusted p-value of less than 0.05. (C) Log2 fold change for genes whose expressions were significantly changed (adjusted p-value <0.05) in *ZNF574* deficient cells under normoxia (y-axis) or hypoxia (x-axis). (D) Bar graph shows the relative fold change in EPO mRNA level under normoxia and hypoxia in *ZNF574* deleted cells. (E) A heatmap of the RNA-sequencing results, describing the difference between wildtype and *ZNF574* deficient HEP3B cells. Red denotes upregulated genes in *ZNF574* deficient cells, blue indicating downregulated genes in *ZNF574* deleted cells. The left column represents gene expression levels under normoxia, and the right column represents gene expression levels under hypoxia.

3.4.8 Discussion

Genome-scale CRISPR screens represent a significant forward-leap in functional genomics, allowing researchers to perturb nearly every gene in the genome in pooled fashion, to identify novel genes that impact a particular phenotype [112]. The high-throughput, unbiased nature of these screens surpasses what is achievable with traditional methods, in which each gene is perturbed independently, thus potentially resulting in an accelerated pace of genetic discovery[107, 111, 129]. A major advantage of CRISPR technology is its specificity; CRISPR/Cas9-mediated gene deletion result in fewer off-target effects than older techniques such as RNA interference (RNAi).

In this report, we describe the development and validation of a HEP3B reporter cell line that serves as a biosensor for EPO production, by inserting eGFP coding sequences in the endogenous locus of *EPO*. To identify new regulators of EPO production, we performed a genome-scale CRISPR knock-out screen using the reporter HEP3B cell line. This screen was performed under normoxia and conditions mimicking hypoxia, to facilitate the identification of genes that potentially impact EPO production independent of HIF.

A notable challenge in CRISPR screens that require FACS sorting of a small proportion of cells, is the depth achieved with these screens. To overcome this limitation, we developed a custom CRISPR library targeting the top 1,255 genes nominated by the genome-scale screen and performed a targeted screen with a significantly higher depth of coverage. This screen identified *ZNF574* as a novel candidate gene that regulates EPO production.

To validate the role of ZNF574 in EPO production, we deleted *ZNF574* in HEP3B cells using 3 independent sgRNAs, and evaluated the impact of this deletion on EPO production. Disruption of *ZNF574* resulted in increased *EPO* mRNA and protein production. Notably, ZNF74 deficiency resulted in increased EPO levels under both normoxia and hypoxia. Additionally, we found that *ZNF574* deletion does not impact most of the genes known to be regulated by HIF. These results suggest that *ZNF574* disruption results in increased *EPO* expression independent of HIF. This finding will be further validated in the future, in a cell line that is deleted for the HIF-responsive elements.

Since sustained HIF activation mediated by PHD inhibitors may result in oncogenesis and neovascularization, targeting *ZNF574* would not be expected to have the same limitations as PHD inhibitors. Future studies aimed at identifying the exact mechanism by which ZNF574 regulates EPO production may result in the identification of novel therapies that promote EPO production independent of HIF.

Chapter 4 Discussion

Erythropoietin (EPO) is a plasma glycoprotein that plays an essential role in regulating red blood cell production[1-8]. Circulating EPO binds its receptor, EPOR, located on the surface of erythroid progenitor cells, resulting in cell survival, proliferation, and differentiation. This regulatory function is critical, as demonstrated by the early embryonic lethality (at ~E12.5) of mice lacking *Epo* or *EpoR* [9, 10]. Although EPO is suspected to have additional roles outside of blood cell formation, its primary and well-established function is to support erythropoiesis [11-16].

Within the hematopoietic system, multipotent stem cells divide and differentiate through several stages to ultimately produce cells that are committed to a specific differentiation fate. The first committed progenitor in the erythroid lineage is known as the 'Burst-forming unit-erythroid' (BFU-E), which develops into 'colony-forming units-erythroid' (CFU-E). CFU-Es then produce proerythroblasts, which differentiate to become basophilic, polychromatic, and orthochromatic erythroblasts sequentially [17, 18]. Orthochromatic erythroblasts discard their nuclei and become reticulocytes, which further mature into red blood cells. While EPO may guide hematopoietic progenitors toward the erythroid lineage, it is dispensable for the very early commitment of multipotent progenitors to the erythroid lineage[3]. The generation of BFU-E and CFU-E is independent of EPO. In contrast, EPO becomes crucial from the CFU-E and proerythroblast stages onward[1-8].

EPO regulation is critical for sustaining an adequate level of erythropoiesis. Overproduction of EPO can lead to erythrocytosis (an abnormal increase in the number of red blood cells), which can in some circumstances increase the risk of thrombosis [86]. Conversely, insufficient EPO production leads to a low red blood cells count (anemia), which when severe results in systemic symptoms, including fatigue and inadequate oxygen transport in the body [82, 83].

The production of EPO is induced under hypoxic conditions, in the liver and kidneys [28-30] [35-38]. This process is regulated by the hypoxia-inducible factor (HIF), which acts as a transcriptional activator for *EPO* [28, 30, 35, 39, 50, 57-64]. In the EPO-producing cells of the liver and kidneys, HIF binds hypoxia-responsive elements at the *EPO* locus, enhancing the transcription of *EPO* [45, 65] [54, 55]. The HIF complex is composed of both an oxygen-sensitive subunit (either HIF1α or HIF2α) and a continuously expressed subunit (HIF1B). The HIF2α subunit is primarily responsible for the upregulation of *EPO* transcription, with a secondary contribution from the HIF1α subunit [68].

The stability of the oxygen-sensitive HIF subunit is closely linked to the availability of oxygen. Under oxygen-rich conditions (normoxia), the HIFA subunit undergoes hydroxylation at specific proline residues by prolyl hydroxylase domain-containing enzymes (PHD1, PHD2, or PHD3). Post-hydroxylation, the modified HIF is recognized by the von Hippel-Lindau protein (VHL), which is part of an E3 ubiquitin ligase complex [69-71]. This leads to the polyubiquitination and subsequent proteasomal degradation of HIFA. Furthermore, the binding of HIF to its transcriptional co-activator p300 is impaired during normoxia because an asparagine residue on HIF is hydroxylated by the enzyme 'factor inhibiting HIF' (FIH), also an oxygen-dependent

process [71, 72]. As a result, the transcription of the *EPO* gene is suppressed during normoxic conditions due to HIF protein degradation and HIF's interaction with p300 is inhibited. In contrast, during hypoxic conditions, the activities of the PHD enzymes and FIH are diminished, allowing HIF to remain stable and to bind the transcriptional co-activator p300, leading to the transcription of HIF-targeted genes, such as *EPO*.

While the role of HIF in EPO regulation is well-established, there is limited understanding of the proteins that regulate EPO secretion and production via HIF-independent mechanisms. Identifying and characterizing such mechanisms could provide insights into alternative regulatory pathways influencing erythropoiesis, potentially revealing new therapeutic targets for treating anemias and other disorders of red blood cell production.

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) screening technology has revolutionized the field of genomics allowing the application of unbiased screens to identify genes involved in various biological processes, including the regulation of erythropoietin (EPO) production and secretion.

As described in chapter two, I performed an unbiased genome-wide loss-of-function screen aimed at identifying proteins that regulate EPO secretion. To perform this screen, I generated a reporter HEK293T cell line that expresses GFP-tagged EPO and, as an internal control, mCherry-tagged A1AT. Through this screen, SURF4 was identified as top candidate gene that impact the intracellular EPO level. Validation studies showed that *SURF4* deletion led to intracellular EPO accumulation and extracellular EPO depletion. In contrast, SURF4

overexpression resulted in the opposite effect. Consistent with the reported localization of SURF4 at the ER membrane, we found an accumulation of EPO in the ER of SURF4 deficient cells. Co-IP experiment showed that SURF4 and EPO physically interact. Collectively, we demonstrate compelling evidence that SURF4 functions as the ER cargo receptor crucial for the efficient secretion of EPO. The studies described above were performed in a reporter heterologous HEK293T cell line. To test the effect of *SURF4* deletion on the secretion of endogenous EPO, we deleted *SURF4* in human HEP3B cells, and confirmed the above findings described in HEK293T.

SURF4 is the mammalian homolog of Erv29p, a yeast protein that aids in the packaging of pro-alpha-factor into COPII vesicles to facilitate its transit from the ER to the Golgi apparatus [117, 125, 126]. Like Erv29p, SURF4 is implicated in similar transport processes in mammalian cells. Currently, only a few known cargos are known to depend on SURF4 for efficient secretion, including APOB, PCSK9, DSSP, AMLEX, and GH [106, 118, 128]. Yet, mice lacking SURF4 die early in embryonic development, hinting at the existence of critical embryonic cargos that are SURF4-dependent [108]. Future studies aimed at identifying the repertoire of proteins that depend on SURF4 for secretion are of high importance in the field.

A recent study introduced the notion that proteins requiring SURF4 for efficient export from the ER share a common "ER-ESCAPE" motif at the beginning of the protein that becomes exposed after the signal sequence is removed [128]. It is this motif that SURF4 supposedly recognizes. However, this high-affinity binding motif for SURF4 is not present in the *EPO* sequence. In contrast, thrombopoietin (TPO), which does possess a N-terminal motif predicted to have

stronger SURF4 affinity, does not depend on SURF4 for secretion. These findings demonstrate that the presence of the "ER-ESCAPE" motif may not be the sole determinant for SURF4's role in the secretion of its cargo proteins.

In the third chapter, we conducted a second genome-scale CRISPR knock-out screen aimed specifically at uncovering regulators of EPO production. For this screen, we developed and validated a HEP3B reporter cell line that expresses eGFP from the endogenous locus of *EPO*. The screen was performed under both normoxic conditions and conditions mimicking hypoxia (using a PHD inhibitor), allowing us to pinpoint genes that may influence EPO production independent of the HIF pathway. This genome-scale screen nominated several genes with a role in EPO production, which were further validated in a secondary screen using a targeted library at a high depth of coverage. This screen identified *ZNF574* as a novel candidate gene that regulates EPO production.

To validate the screen findings, we deleted *ZNF574* in HEP3B cells and found that ZNF574 deficiency results in increased *EPO* mRNA and protein production, both under normoxia and hypoxia. Moreover, deletion of *ZNF574* appeared to have no noticeable impact on the expression of the majority of genes that are known to be regulated by HIF. These findings suggest that *ZNF574* disruption results in increased EPO production independently of HIF.

As demonstrated in this dissertation, utilizing unbiased genome-scale CRISPR screens, I discovered SURF4 and ZNF575 as two previously unidentified regulators of EPO secretion and EPO production, respectively. Further studies are warranted to elucidate the pathways and

mechanisms by which ZNF574 (and additional proteins uncovered in the CRISPR screens) regulate EPO production.

Bibliography

1. Gregory, C.J., *Erythropoietin sensitivity as a differentiation marker in the hemopoietic system: studies of three erythropoietic colony responses in culture.* J Cell Physiol, 1976. **89**(2): p. 289-301.
2. Suzuki, N., et al., *Erythroid-specific expression of the erythropoietin receptor rescued its null mutant mice from lethality.* Blood, 2002. **100**(7): p. 2279-88.
3. Wu, H., et al., *Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor.* Cell, 1995. **83**(1): p. 59-67.
4. Hidalgo, D., et al., *EpoR stimulates rapid cycling and larger red cells during mouse and human erythropoiesis.* Nat Commun, 2021. **12**(1): p. 7334.
5. Koury, M.J. and M.C. Bondurant, *Erythropoietin retards DNA breakdown and prevents programmed death in erythroid progenitor cells.* Science, 1990. **248**(4953): p. 378-81.
6. Koury, M.J., M.C. Bondurant, and J.B. Atkinson, *Erythropoietin control of terminal erythroid differentiation: maintenance of cell viability, production of hemoglobin, and development of the erythrocyte membrane.* Blood Cells, 1987. **13**(1-2): p. 217-26.
7. Liu, Y., et al., *Suppression of Fas-FasL coexpression by erythropoietin mediates erythroblast expansion during the erythropoietic stress response in vivo.* Blood, 2006. **108**(1): p. 123-33.
8. Spivak, J.L., et al., *Erythropoietin is both a mitogen and a survival factor.* Blood, 1991. **77**(6): p. 1228-33.
9. Yamazaki, S., et al., *A mouse model of adult-onset anaemia due to erythropoietin deficiency.* Nat Commun, 2013. **4**: p. 1950.
10. Zeigler, B.M., et al., *A mouse model for an erythropoietin-deficiency anemia.* Dis Model Mech, 2010. **3**(11-12): p. 763-72.
11. Arcasoy, M.O., *The non-haematopoietic biological effects of erythropoietin.* Br J Haematol, 2008. **141**(1): p. 14-31.
12. Jelkmann, W., *Erythropoietin: back to basics.* Blood, 2010. **115**(21): p. 4151-2.
13. Morita, M., et al., *HLF/HIF-2alpha is a key factor in retinopathy of prematurity in association with erythropoietin.* EMBO J, 2003. **22**(5): p. 1134-46.
14. Noguchi, C.T., et al., *Survival and proliferative roles of erythropoietin beyond the erythroid lineage.* Expert Rev Mol Med, 2008. **10**: p. e36.
15. Scheerer, N., et al., *The anemia of the newborn induces erythropoietin expression in the developing mouse retina.* Am J Physiol Regul Integr Comp Physiol, 2010. **299**(1): p. R111-8.
16. Teng, R., et al., *Disrupted erythropoietin signalling promotes obesity and alters hypothalamus proopiomelanocortin production.* Nat Commun, 2011. **2**: p. 520.
17. Grover, A., et al., *Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate.* J Exp Med, 2014. **211**(2): p. 181-8.

18. Tusi, B.K., et al., *Population snapshots predict early haematopoietic and erythroid hierarchies.* Nature, 2018. **555**(7694): p. 54-60.
19. Cheetham, J.C., et al., *NMR structure of human erythropoietin and a comparison with its receptor bound conformation.* Nat Struct Biol, 1998. **5**(10): p. 861-6.
20. Syed, R.S., et al., *Efficiency of signalling through cytokine receptors depends critically on receptor orientation.* Nature, 1998. **395**(6701): p. 511-6.
21. Elliott, S., et al., *Enhancement of therapeutic protein in vivo activities through glycoengineering.* Nat Biotechnol, 2003. **21**(4): p. 414-21.
22. Carnot, P. and C. Deflandre, *The hemopoietic activity of serum during the regeneration of blood.* Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences, 1906. **143**: p. 384-386.
23. Reissmann, K.R., *Studies on the mechanism of erythropoietic stimulation in parabiotic rats during hypoxia.* Blood, 1950. **5**(4): p. 372-80.
24. Erslev, A., *Humoral regulation of red cell production.* Blood, 1953. **8**(4): p. 349-57.
25. Lin, F.K., et al., *Cloning and expression of the human erythropoietin gene.* Proc Natl Acad Sci U S A, 1985. **82**(22): p. 7580-4.
26. Jacobs, K., et al., *Isolation and characterization of genomic and cDNA clones of human erythropoietin.* Nature, 1985. **313**(6005): p. 806-10.
27. Lee-Huang, S., *Cloning and expression of human erythropoietin cDNA in Escherichia coli.* Proc Natl Acad Sci U S A, 1984. **81**(9): p. 2708-12.
28. Blumberg, A., H. Keller, and H.R. Marti, *Effect of altitude on erythropoiesis and oxygen affinity in anaemic patients on maintenance dialysis.* Eur J Clin Invest, 1973. **3**(2): p. 93-7.
29. Naets, J.P., *[The role of the kidney in erythropoiesis].* Acta Clin Belg, 1960. **15**: p. 361-496.
30. Nathan, D.G., et al., *Erythropoiesis in Anephric Man.* J Clin Invest, 1964. **43**(11): p. 2158-65.
31. Bachmann, S., M. Le Hir, and K.U. Eckardt, *Co-localization of erythropoietin mRNA and ecto-5'-nucleotidase immunoreactivity in peritubular cells of rat renal cortex indicates that fibroblasts produce erythropoietin.* J Histochem Cytochem, 1993. **41**(3): p. 335-41.
32. Koury, S.T., et al., *Quantitation of erythropoietin-producing cells in kidneys of mice by in situ hybridization: correlation with hematocrit, renal erythropoietin mRNA, and serum erythropoietin concentration.* Blood, 1989. **74**(2): p. 645-51.
33. Maxwell, P.H., et al., *Sites of erythropoietin production.* Kidney Int, 1997. **51**(2): p. 393-401.
34. Maxwell, P.H., et al., *Identification of the renal erythropoietin-producing cells using transgenic mice.* Kidney Int, 1993. **44**(5): p. 1149-62.
35. Eckardt, K.U., et al., *Age-dependent expression of the erythropoietin gene in rat liver and kidneys.* J Clin Invest, 1992. **89**(3): p. 753-60.
36. Jacobson, L.O., et al., *Role of the kidney in erythropoiesis.* Nature, 1957. **179**(4560): p. 633-4.
37. Schooley, J.C. and L.J. Mahlmann, *Erythropoietin production in the anephric rat. I. Relationship between nephrectomy, time of hypoxic exposure, and erythropoietin production.* Blood, 1972. **39**(1): p. 31-8.
38. Zanjani, E.D., et al., *Studies on the liver to kidney switch of erythropoietin production.* J Clin Invest, 1981. **67**(4): p. 1183-8.

39. Fried, W., *The liver as a source of extrarenal erythropoietin production.* Blood, 1972. **40**(5): p. 671-7.
40. Suzuki, N., et al., *Erythropoietin production in neuroepithelial and neural crest cells during primitive erythropoiesis.* Nat Commun, 2013. **4**: p. 2902.
41. Zanjani, E.D., et al., *Liver as the primary site of erythropoietin formation in the fetus.* J Lab Clin Med, 1977. **89**(3): p. 640-4.
42. Carmena, A.O., D. Howard, and F. Stohlman, Jr., *Regulation of erythropoiesis. XXII. Erythropoietin production in the newborn animal.* Blood, 1968. **32**(3): p. 376-82.
43. Lucarelli, G., et al., *Fetal and neonatal erythropoiesis.* Ann N Y Acad Sci, 1968. **149**(1): p. 544-59.
44. Peschle, C., et al., *Erythropoietin production by the liver in fetal-neonatal life.* Life Sci, 1975. **17**(8): p. 1325-30.
45. Suzuki, N., et al., *Specific contribution of the erythropoietin gene 3' enhancer to hepatic erythropoiesis after late embryonic stages.* Mol Cell Biol, 2011. **31**(18): p. 3896-905.
46. Zanjani, E.D., et al., *Erythropoietin production in the fetus: role of the kidney and maternal anemia.* J Lab Clin Med, 1974. **83**(2): p. 281-7.
47. Lucarelli, G., D. Howard, and F. Stohlman, Jr., *Regulation of Erythropoiesis. Xv. Neonatal Erythropoiesis and the Effect of Nephrectomy.* J Clin Invest, 1964. **43**(11): p. 2195-203.
48. Koury, S.T., et al., *Localization of cells producing erythropoietin in murine liver by in situ hybridization.* Blood, 1991. **77**(11): p. 2497-503.
49. Maxwell, P.H., et al., *Expression of a homologously recombined erythopoietin-SV40 T antigen fusion gene in mouse liver: evidence for erythropoietin production by Ito cells.* Blood, 1994. **84**(6): p. 1823-30.
50. Flake, A.W., et al., *Erythropoietin production by the fetal liver in an adult environment.* Blood, 1987. **70**(2): p. 542-5.
51. Dame, C., et al., *Hepatic erythropoietin gene regulation by GATA-4.* J Biol Chem, 2004. **279**(4): p. 2955-61.
52. Gupta, M., P.T. Mungai, and E. Goldwasser, *A new transacting factor that modulates hypoxia-induced expression of the erythropoietin gene.* Blood, 2000. **96**(2): p. 491-7.
53. Makita, T., et al., *A developmental transition in definitive erythropoiesis: erythropoietin expression is sequentially regulated by retinoic acid receptors and HNF4.* Genes Dev, 2001. **15**(7): p. 889-901.
54. Semenza, G.L., et al., *Cell-type-specific and hypoxia-inducible expression of the human erythropoietin gene in transgenic mice.* Proc Natl Acad Sci U S A, 1991. **88**(19): p. 8725-9.
55. Hirano, I., et al., *Renal Anemia Model Mouse Established by Transgenic Rescue with an Erythropoietin Gene Lacking Kidney-Specific Regulatory Elements.* Mol Cell Biol, 2017. **37**(4).
56. Madan, A., et al., *Regulated basal, inducible, and tissue-specific human erythropoietin gene expression in transgenic mice requires multiple cis DNA sequences.* Blood, 1995. **85**(10): p. 2735-41.
57. Bernhardt, W.M., et al., *Inhibition of prolyl hydroxylases increases erythropoietin production in ESRD.* J Am Soc Nephrol, 2010. **21**(12): p. 2151-6.

58. DeGowin, R.L., et al., *Erythropoiesis and erythropoietin in patients with chronic renal failure treated with hemodialysis and testosterone.* Ann Intern Med, 1970. **72**(6): p. 913-8.
59. Erslev, A.J., *Erythropoietic function in uremic rabbits.* AMA Arch Intern Med, 1958. **101**(2): p. 407-17.
60. Gallagher, N.I., C.J. Mc, and R.D. Lange, *Erythropoietin production in uremic rabbits.* J Lab Clin Med, 1961. **57**: p. 281-9.
61. Mirand, E.A., et al., *Extra-renal production of erythropoietin in man.* Acta Haematol, 1968. **39**(6): p. 359-65.
62. Mirand, E.A. and T.C. Prentice, *Presence of plasma erythropoietin in hypoxic rats with or without kidney (s) and/or spleen.* Proc Soc Exp Biol Med, 1957. **96**(1): p. 49-51.
63. Naets, J.P. and M. Wittek, *Presence of erythropoietin in the plasma of one anephric patient.* Blood, 1968. **31**(2): p. 249-51.
64. Tan, C.C., K.U. Eckardt, and P.J. Ratcliffe, *Organ distribution of erythropoietin messenger RNA in normal and uremic rats.* Kidney Int, 1991. **40**(1): p. 69-76.
65. Semenza, G.L., et al., *Hypoxia-inducible nuclear factors bind to an enhancer element located 3′ to the human erythropoietin gene.* Proc Natl Acad Sci U S A, 1991. **88**(13): p. 5680-4.
66. Beck, I., et al., *Enhancer element at the 3′-flanking region controls transcriptional response to hypoxia in the human erythropoietin gene.* J Biol Chem, 1991. **266**(24): p. 15563-6.
67. Pugh, C.W., et al., *Functional analysis of an oxygen-regulated transcriptional enhancer lying 3′ to the mouse erythropoietin gene.* Proc Natl Acad Sci U S A, 1991. **88**(23): p. 10553-7.
68. Rosenberger, C., et al., *Expression of hypoxia-inducible factor-1alpha and -2alpha in hypoxic and ischemic rat kidneys.* J Am Soc Nephrol, 2002. **13**(7): p. 1721-32.
69. Bunn, H.F., *Erythropoietin.* Cold Spring Harb Perspect Med, 2013. **3**(3): p. a011619.
70. Haase, V.H., *Regulation of erythropoiesis by hypoxia-inducible factors.* Blood Rev, 2013. **27**(1): p. 41-53.
71. Schofield, C.J. and P.J. Ratcliffe, *Oxygen sensing by HIF hydroxylases.* Nat Rev Mol Cell Biol, 2004. **5**(5): p. 343-54.
72. Mahon, P.C., K. Hirota, and G.L. Semenza, *FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity.* Genes Dev, 2001. **15**(20): p. 2675-86.
73. Imagawa, S., M. Yamamoto, and Y. Miura, *GATA transcription factors negatively regulate erythropoietin gene expression.* Acta Haematol, 1996. **95**(3-4): p. 248-56.
74. Imagawa, S., T. Izumi, and Y. Miura, *Positive and negative regulation of the erythropoietin gene.* J Biol Chem, 1994. **269**(12): p. 9038-44.
75. Obara, N., et al., *Repression via the GATA box is essential for tissue-specific erythropoietin gene expression.* Blood, 2008. **111**(10): p. 5223-32.
76. Dame, C., et al., *Wilms tumor suppressor, Wt1, is a transcriptional activator of the erythropoietin gene.* Blood, 2006. **107**(11): p. 4282-90.
77. Sanchez-Elsner, T., et al., *A cross-talk between hypoxia and TGF-beta orchestrates erythropoietin gene regulation through SP1 and Smads.* J Mol Biol, 2004. **336**(1): p. 9-24.

78. Ji, F., et al., *Brain-specific Wt1 deletion leads to depressive-like behaviors in mice via the recruitment of Tet2 to modulate Epo expression.* Mol Psychiatry, 2021. **26**(8): p. 4221-4233.
79. Galson, D.L., et al., *The orphan receptor hepatic nuclear factor 4 functions as a transcriptional activator for tissue-specific and hypoxia-specific erythropoietin gene expression and is antagonized by EAR3/COUP-TF1.* Mol Cell Biol, 1995. **15**(4): p. 2135-44.
80. da Cunha, M.S., et al., *Vitamin A deficiency modulates iron metabolism via ineffective erythropoiesis.* J Nutr Biochem, 2014. **25**(10): p. 1035-44.
81. Zmajkovic, J., et al., *A Gain-of-Function Mutation in EPO in Familial Erythrocytosis.* N Engl J Med, 2018. **378**(10): p. 924-930.
82. Kim, A.R., et al., *Functional Selectivity in Cytokine Signaling Revealed Through a Pathogenic EPO Mutation.* Cell, 2017. **168**(6): p. 1053-1064 e15.
83. Babitt, J.L. and H.Y. Lin, *Mechanisms of anemia in CKD.* J Am Soc Nephrol, 2012. **23**(10): p. 1631-4.
84. Hellstrom-Lindberg, E., *Efficacy of erythropoietin in the myelodysplastic syndromes: a meta-analysis of 205 patients from 17 studies.* Br J Haematol, 1995. **89**(1): p. 67-71.
85. Jadersten, M., et al., *Erythropoietin and granulocyte-colony stimulating factor treatment associated with improved survival in myelodysplastic syndrome.* J Clin Oncol, 2008. **26**(21): p. 3607-13.
86. Patnaik, M.M. and A. Tefferi, *The complete evaluation of erythrocytosis: congenital and acquired.* Leukemia, 2009. **23**(5): p. 834-44.
87. Braakman, I. and N.J. Bulleid, *Protein Folding and Modification in the Mammalian Endoplasmic Reticulum.* Annual Review of Biochemistry, 2011. **80**(1): p. 71-99.
88. Mathias Uhlen, L.F., Bjorn M. Hallstrom,Cecilia Lindskog, Per Oksvold, Adil Mardinoglu, Åsa Sivertsson, Caroline Kampf, Evelina Sjöstedt, Anna Asplund, IngMarie Olsson, Karolina Edlund, Emma Lundberg, Sanjay Navani,, et al., *Tissue-based map of the human proteome.* Science, 2015. **347**(6220).
89. Jensen, D. and R. Schekman, *COPII-mediated vesicle formation at a glance.* J Cell Sci, 2011. **124**(Pt 1): p. 1-4.
90. Barlowe, C. and A. Helenius, *Cargo Capture and Bulk Flow in the Early Secretory Pathway.* Annu Rev Cell Dev Biol, 2016. **32**: p. 197-222.
91. Moise Bendayan, et al., *Quantitative immunocytochemical localization of pancreatic secretory proteins in subcellular compartments of the rat acinar cell.* The Journal of Histochemistry and Cytochemistry, 1980. **28**(2): p. 149-160.
92. Kuehn, M.J., J.M. Herrmann, and R. Schekman, *COPII-cargo interactions direct protein sorting into ER-derived transport vesicles.* Nature, 1998. **391**(8): p. 187-190.
93. Nina R.Salama, T.Y., Randy W.Schekman, *The Sec13p complex and reconstitution of vesicle budding from the ER with purified cytosolic proteins.* The EMBO Journal, 1993. **12**(11): p. 4073-4082.
94. Balch, W.E., et al., *Vesicular stomatitis virus glycoprotein is sorted and concentrated during export from the endoplasmic reticulum.* Cell, 1994. **76**: p. 841-852.
95. Charles Barlowe, et al., *COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum.* Cell, 1994. **77**: p. 895-907.
96. Kappeler, F., et al., *The recycling of ERGIC-53 in the Early Secretory Pathway.* The Journal of Biological Chemistry, 1997. **272**(50): p. 31801-31808.

97. Itin, C., et al., *ERGIC-53 Is a Functional Mannose-selective and Calcium-dependent Human Homologue of Leguminous Lectins.* Molecular Biology of the Cell, 1996. **7**: p. 483-493.
98. William C. Nichols, et al., *Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII.* Cell, 1998. **93**.
99. Zhang, B., et al., *Mice deficient in LMAN1 exhibit FV and FVIII deficiencies and liver accumulation of alpha1-antitrypsin.* Blood, 2011. **118**(12): p. 3384-91.
100. Zhang, B., *Recent developments in the understanding of the combined deficiency of FV and FVIII.* Br J Haematol, 2009. **145**(1): p. 15-23.
101. Christian Appenzeller, et al., *The lectin ERGIC-53 is a cargo transport receptor for glycoproteins.* Nature Cell Biology, 1999. **1**.
102. Nyfeler, B., et al., *Identification of ERGIC-53 as an intracellular transport receptor of alpha1-antitrypsin.* J Cell Biol, 2008. **180**(4): p. 705-12.
103. Gomez-Navarro, N. and E. Miller, *Protein sorting at the ER–Golgi interface.* The Journal of Cell Biology, 2016. **215**(6): p. 769-778.
104. Kim, W.Y., et al., *Failure to prolyl hydroxylate hypoxia-inducible factor alpha phenocopies VHL inactivation in vivo.* The EMBO Journal, 2006. **25**.
105. Kochling, J., P.T. Curtin, and A. Madan, *Regulation of human erythropoietin gene induction by upstream flanking sequences in transgenic mice.* British Journal of Haematology, 1998. **103**(4).
106. Emmer, B.T., et al., *The cargo receptor SURF4 promotes the efficient cellular secretion of PCSK9.* Elife, 2018. **7**.
107. Sanjana, N.E., O. Shalem, and F. Zhang, *Improved vectors and genome-wide libraries for CRISPR screening.* Nat Methods, 2014. **11**(8): p. 783-784.
108. Emmer, B.T., et al., *Murine Surf4 is essential for early embryonic development.* BioRxiv, 2019.
109. Wei Li, et al., *MAGeCK enables robust identification of essential genes from genome-scale CRISPR:Cas9 knockout screens.* Genome Biology, 2014. **15**(554).
110. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
111. Joung, J., et al., *Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening.* Nat Protoc, 2017. **12**(4): p. 828-863.
112. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.
113. Khoriaty, R., et al., *Absence of a red blood cell phenotype in mice with hematopoietic deficiency of SEC23B.* Mol Cell Biol, 2014. **34**(19): p. 3721-34.
114. Khoriaty, R., et al., *Functions of the COPII gene paralogs SEC23A and SEC23B are interchangeable in vivo.* Proc Natl Acad Sci U S A, 2018. **115**(33): p. E7748-E7757.
115. Costantini, L.M., et al., *A palette of fluorescent proteins optimized for diverse cellular environments.* Nat Commun, 2015. **6**: p. 7670.
116. Luker, K.E., et al., *Comparative study reveals better far-red fluorescent protein for whole body imaging.* Sci Rep, 2015. **5**: p. 10332.
117. Belden, W.J. and a.C. Barlowe, *Role of Erv29p in Collecting Soluble Secretory Proteins into ER-Derived Transport Vesicles.* Science, 2001. **294**(5546): p. 1528-1531.

118. Saegusa, K., et al., *SFT-4/Surf4 control ER export of soluble cargo proteins and participate in ER exit site organization.* The Journal of Cell Biology, 2018. **217**(6): p. 2073-2085.
119. Falck, D., et al., *Affinity purification of erythropoietin from cell culture supernatant combined with MALDI-TOF-MS analysis of erythropoietin N-glycosylation.* Sci Rep, 2017. **7**(1): p. 5324.
120. Goldwasser, E., C.K. Kung, and J. Eliason, *On the mechanism of erythropoietin-induced differentiation. 13. The role of sialic acid in erythropoietin action.* J Biol Chem, 1974. **249**(13): p. 4202-6.
121. Lai, P.H., et al., *Structural characterization of human erythropoietin.* J Biol Chem, 1986. **261**(7): p. 3116-21.
122. Maley, F., et al., *Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases.* Anal Biochem, 1989. **180**(2): p. 195-204.
123. Trumbly, R.J., et al., *Amplified expression of streptomyces endo-beta-N-acetylglucosaminidase H in Escherichia coli and characterization of the enzyme product.* J Biol Chem, 1985. **260**(9): p. 5683-90.
124. Sandra Mitrovic, et al., *The Cargo Receptors Surf4, Endoplasmic Reticulum-Golgi Intermediate Compartment (ERGIC)-53, and p25 Are Required to Maintain the Architecture of ERGIC and Golgi.* Molecular Biology of the Cell, 2008. **19**: p. 1976-1990.
125. Malkus, P., F. Jiang, and R. Schekman, *Concentrative sorting of secretory cargo proteins into COPII-coated vesicles.* J Cell Biol, 2002. **159**(6): p. 915-21.
126. Otte, S. and C. Barlowe, *Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles.* Nat Cell Biol, 2004. **6**(12): p. 1189-94.
127. Foley, D.A., H.J. Sharpe, and S. Otte, *Membrane topology of the endoplasmic reticulum to Golgi transport factor Erv29p.* Mol Membr Biol, 2007. **24**(4): p. 259-68.
128. Yin, Y., et al., *Surf4 (Erv29p) binds amino-terminal tripeptide motifs of soluble cargo proteins with different affinities, enabling prioritization of their exit from the endoplasmic reticulum.* PLoS Biol, 2018. **16**(8): p. e2005140.
129. Ihry, R.J., et al., *Genome-Scale CRISPR Screens Identify Human Pluripotency-Specific Genes.* Cell Rep, 2019. **27**(2): p. 616-630 e6.
130. Canver, M.C., et al., *BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis.* Nature, 2015. **527**(7577): p. 192-7.
131. Yamauchi, T., et al., *Genome-wide CRISPR-Cas9 Screen Identifies Leukemia-Specific Dependence on a Pre-mRNA Metabolic Pathway Regulated by DCPS.* Cancer Cell, 2018. **33**(3): p. 386-400 e5.
132. Guang L. Wang, et al., *Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular 02 tension.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**: p. 5510-5514.
133. J.D. Firth, et al., *Oxygen-regulated control elements in the phosphoglycerate kinase 1 and lactate dehydrogenase A genes: similarities with the erythropoietin 3' enhancer.* Proc Natl Acad Sci U S A, 1994. **91**: p. 6496-6500.
134. Franke, K., M. Gassmann, and B. Wielockx, *Erythrocytosis: the HIF pathway in control.* Blood, 2013. **122**(7): p. 1122-8.
135. Kapitsinou, P.P., et al., *Hepatic HIF-2 regulates erythropoietic responses to hypoxia in renal anemia.* Blood, 2010. **116**(16): p. 3039-48.

136. Ladroue, C., et al., *Distinct deregulation of the hypoxia inducible factor by PHD2 mutants identified in germline DNA of patients with polycythemia.* Haematologica, 2012. **97**(1): p. 9-14.
137. Percy, M.J., et al., *Two new mutations in the HIF2A gene associated with erythrocytosis.* Am J Hematol, 2012. **87**(4): p. 439-42.
138. Zhuang, Z., et al., *Somatic HIF2A gain-of-function mutations in paraganglioma with polycythemia.* N Engl J Med, 2012. **367**(10): p. 922-30.
139. Gruber, M., et al., *Acute postnatal ablation of Hif-2alpha results in anemia.* Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2301-6.
140. Rankin, E.B., et al., *Inactivation of the arylhydrocarbon receptor nuclear translocator (Arnt) suppresses von Hippel-Lindau disease-associated vascular tumors in mice.* Mol Cell Biol, 2005. **25**(8): p. 3163-72.
141. Sankaran, V.G. and M.J. Weiss, *Anemia: progress in molecular mechanisms and therapies.* Nat Med, 2015. **21**(3): p. 221-30.
142. Joharapurkar, A.A., et al., *Prolyl Hydroxylase Inhibitors: A Breakthrough in the Therapy of Anemia Associated with Chronic Diseases.* J Med Chem, 2018. **61**(16): p. 6964-6982.
143. Maxwell, P.H. and K.U. Eckardt, *HIF prolyl hydroxylase inhibitors for the treatment of renal anaemia and beyond.* Nat Rev Nephrol, 2016. **12**(3): p. 157-68.
144. Kaplan, J., *Roxadustat and Anemia of Chronic Kidney Disease.* the New England Journal of Medicine, 2019.
145. Gupta, N. and J.B. Wish, *Hypoxia-Inducible Factor Prolyl Hydroxylase Inhibitors: A Potential New Treatment for Anemia in Patients With CKD.* Am J Kidney Dis, 2017. **69**(6): p. 815-826.
146. Yamamoto, A., et al., *Systemic silencing of PHD2 causes reversible immune regulatory dysfunction.* J Clin Invest, 2019. **130**.
147. Semenza, G.L., *Targeting HIF-1 for cancer therapy.* Nat Rev Cancer, 2003. **3**(10): p. 721-32.
148. Masoud, G.N. and W. Li, *HIF-1alpha pathway: role, regulation and intervention for cancer therapy.* Acta Pharm Sin B, 2015. **5**(5): p. 378-89.
149. Lily Jun-shen Huang, Stefan N. Constantinescu, and a.H.F. Lodish, *The N-Terminal Domain of Janus Kinase 2Is Required for Golgi Processing and Cell SurfaceExpression of Erythropoietin Receptor.* Molecular Cell, 2001. **8**: p. 1327-1338.
150. Halupa, A., et al., *A novel role for STAT1 in regulating murine erythropoiesis: deletion of STAT1 results in overall reduction of erythroid progenitors and alters their distribution.* Blood, 2005. **105**(2): p. 552-61.
151. Victor R. Gordeuk, David W. Stockton, and J.T. Prchal, *Congenital Polycythemias/Erythrocytoses.* Haematologica, 2005.
152. T Ng, et al., *Recombinant erythropoietin in clinical practice.* BMJ Journals, 2003. **79**: p. 367-376.
153. J. Douglas Rizzo, et al., *Erythropoietin: A Paradigm for the Development of Practice Guidelines.* ASH Education book, 2001.
154. *Erythropoietin (EPO) Drugs Market Analysis By Drug Class (Biologics, Biosimilars), By Product (Epoetin-alfa, Epoetin-beta, Darbepoetin-alfa, Others), By Application, And Segment Forecasts, 2018 - 2025*. [blog] 2017 May, 2017; Available from: https://www.grandviewresearch.com/industry-analysis/erythropoietin-epo-drugs-market.

155. McKoy, J.M., et al., *Epoetin-associated pure red cell aplasia: past, present, and future considerations.* Transfusion, 2008. **48**(8): p. 1754-62.
156. Mastromarino, V., et al., *Erythropoietin in cardiac disease: effective or harmful?* J Cardiovasc Med (Hagerstown), 2013. **14**(12): p. 870-8.
157. Lippi, G., M. Franchini, and E.J. Favaloro, *Thrombotic complications of erythropoiesis-stimulating agents.* Semin Thromb Hemost, 2010. **36**(5): p. 537-49.
158. Bonomini, M., et al., *New Treatment Approaches for the Anemia of CKD.* Am J Kidney Dis, 2016. **67**(1): p. 133-42.
159. Locatelli, F., et al., *Kidney Disease: Improving Global Outcomes guidelines on anaemia management in chronic kidney disease: a European Renal Best Practice position statement.* Nephrol Dial Transplant, 2013. **28**(6): p. 1346-59.
160. Macdougall, I.C. and A.C. Cooper, *Erythropoietin resistance: the role of inflammation and pro-inflammatory cytokines.* Nephrol Dial Transplant, 2002. **17 Suppl 11**: p. 39-43.
161. Shimizu, H., et al., *Pure red cell aplasia induced only by intravenous administration of recombinant human erythropoietin.* Acta Haematol, 2011. **126**(2): p. 114-8.
162. Kurata, Y., T. Tanaka, and M. Nangaku, *An evaluation of roxadustat for the treatment of anemia associated with chronic kidney disease.* Expert Opin Pharmacother, 2022. **23**(1): p. 19-28.
163. Zhang, Y., et al., *HIF-1alpha is necessary for activation and tumour-promotion effect of cancer-associated fibroblasts in lung cancer.* J Cell Mol Med, 2021. **25**(12): p. 5457-5469.
164. Maxwell, P.H., *The HIF pathway in cancer.* Semin Cell Dev Biol, 2005. **16**(4-5): p. 523-30.
165. Singhal, R., et al., *HIF-2alpha activation potentiates oxidative cell death in colorectal cancers by increasing cellular iron.* J Clin Invest, 2021. **131**(12).
166. Wicks, E.E. and G.L. Semenza, *Hypoxia-inducible factors: cancer progression and clinical translation.* J Clin Invest, 2022. **132**(11).
167. Pezzuto, A. and E. Carico, *Role of HIF-1 in Cancer Progression: Novel Insights. A Review.* Curr Mol Med, 2018. **18**(6): p. 343-351.
168. Lin, Z., et al., *The Endoplasmic Reticulum Cargo Receptor SURF4 Facilitates Efficient Erythropoietin Secretion.* Mol Cell Biol, 2020. **40**(23).