Genetic studies of hereditary myeloproliferative disorders

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von
Jakub Zmajkovič
aus Bratislava, Slowakei

Basel, 2019

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Professor Dr. med Radek C. Skoda

Professor Dr. Mihaela Zavolan

Basel, den 11.12. 2018

Prof. Dr. Martin Spiess
Dekan der Philosophisch-
Naturwissenschaftlichen Fakultät

Table of Contents

Acknowledgments

I would like to sincerely thank everybody that contributed to this dissertation in any manner.

Radek, thank you very much for giving me the opportunity to join your lab and work on several exciting projects. I have learned a lot in terms of scientific excellence and also human relationships. However, what I value the most is 1. a scientific rigor you use to assess your work, and you also demand from the people in the lab (I will never forget Sunday evenings with you aligning the axes in the figures) and 2. independence, I gained under your supervision. All of above will be invaluable in my future career and I am very grateful for it.

The same gratitude goes to my colleagues and friends in the lab: **Morgane, Ronny, Takafumi, Pontus, Jean, Basia, Lucia, Yukiko, Julian, Shivam** and **Nils** for the warm welcome and the great time we spent together. You were the best lab mates one can imagine, and I hope to stay in touch with you also after we move to other places. Special thank you also goes to our good souls in the lab, **Gabi, Hélène** and **Hui.** Thank you very much for your everyday help running the lab smoothly and supporting us immensely!

I am also thankful to **Mihaela** and **Lukas** – thank you for your support and guidance in the committee meetings.

Many thanks go as well to the colleagues from the other hematology labs (Myeloid malignancies, Childhood leukemia, and Stem cells and hematopoiesis) for their support! It was great to share our lab space with you, as it led to unanticipated and fruitful discussions about science and everyday life. Furthermore, I would like to thank all the core facilities members for providing us with the expertise and framework to perform our experiments and Swiss Cancer League, Swiss National Foundation and the University of Basel for supporting my research.

Nakoniec by som rád poďakoval mojim najbližším, ktorí počas môjho doktorandského štúdia stáli pri mne a bez ktorých by táto práca určite nevznikla. Ďakujem **Katka**, za tvoju každodennú starostlivosť o nás, vždy si mi bola a budeš mojou vedeckou a nevedeckou oporou. Ďakujem aj nášmu slniečku – **Ninka**, vďaka tebe sa mi vždy podarilo rýchlo zabudnúť na pracovné povinnosti a problémy, tvoj smiech je pre mňa balzam na dušu. Ďakujem **mami** a **tati**, za všetku vašu podporu od začiatku až doteraz, nikdy vám to nezabudnem. Ďakujem aj všetkým ostatným členom mojej rodiny a všetkým kamarátom, za vašu lásku, pomoc a podporu.

Summary

More than 300 billion blood cells are being replaced daily in a process called hematopoiesis. Hematopoiesis is orchestrated by hematopoietic stem cells (HSCs) residing in the bone marrow. HSCs produce multipotent and lineage-restricted progenitors, that are responsible for the supply of mature blood cells. Production of blood cells is governed by hematopoietic growth factors that are required for the survival and proliferation of blood cells at all stages of development. Mutations in genes responsible for the regulation of this fine-tuned system cause aberrant proliferation of different blood compartments. Myeloproliferative neoplasms (MPN) are characterized by the abnormal expansion of erythroid, megakaryocytic, and myeloid lineages, that is caused either by somatic mutations or by germline mutations transmitted through Mendelian inheritance within the family. The main topic of my doctoral research was the investigation of two distinct pedigrees diagnosed with erythrocytosis and thrombocytosis, respectively.

Erythrocytosis occurred in ten individuals of Norwegian family that presented elevated hemoglobin and erythropoietin (EPO) serum levels. We performed genome-wide linkage analysis using SNP arrays coupled with targeted sequencing and identified a heterozygous single base deletion (ΔG) in exon 2 of the *EPO* gene as the sole candidate gene mutation in affected family members. EPO stimulates the proliferation of erythrocyte progenitors and prevents their apoptosis in order to produce mature erythrocytes. Surprisingly, ΔG introduces a frame-shift that generates a novel, 51-residue long polypeptide, which would predict a loss of erythropoietin function, and is at odds with the erythrocytosis phenotype. To elucidate the mechanism by which the loss-of-function mutation causes gain-of-function phenotype, we utilized the CRISPR/Cas9 genome editing to introduce the ΔG mutation into Hep3B cells, a human hepatoma cell line that expresses EPO. We found that cells with ΔG mutation produce excessive amounts of biologically active EPO and reproduces the observation form the affected family members. On the molecular level, in addition to the known transcript originating from the physiologic promoter (P1), we identified novel transcripts that initiate in intron 1 of *EPO* from a putative alternative promoter (P2). Further functional analysis of P2 mRNAs revealed an alternative translational start site in exon 2 that P2 transcripts use to produce a biologically active EPO protein, by fusing a novel N-terminus to the EPO

coding sequence through the ΔG single base deletion. Our data demonstrate for the first time, that a mutation in *EPO* cause familial erythrocytosis and explain how the ΔG mutation results in a gain-function phenotype.

I also investigated a pedigree with autosomal-dominant. Targeted sequencing identified a novel activating mutation in exon 3 of the thrombopoietin (*THPO)* gene, a single nucleotide G->T substitution. Thrombopoietin stimulates the production of platelets from megakaryocytes. *THPO* expression is regulated on the translational level by seven upstream open reading frames (uORF1-7) in the exons 1-3 of *THPO* mRNA, that are interfering with the translation of TPO. G>T mutation maps to the Kozak sequence of the uORF7, the most critical negative regulator of TPO translation. We performed TPO overexpression and *in vitro* translation experiments to demonstrate that the G>T mutation disrupts the negative regulation governed by uORF7 and allows for increased translation of *THPO* protein coding sequence, ultimately causing thrombocytosis.

Collectively, in both studies we identified novel gain-of-function mutations in hematopoietic growth factors, that act at different steps of gene expression and result in the dysregulated production of EPO and TPO, causing erythrocytosis and thrombocytosis in respective pedigrees.

1 Introduction

1.1 Hematopoiesis

The human body is a complex organism composed of 10^{13} to 10^{14} human cells. Strikingly, as illustrated in Figure 1, around 93% of cells are coming from a single tissue – blood (Sender et al., 2016). This tissue is composed of more than ten different cell types that perform various functions, ranging from oxygen and nutrient transport to immune surveillance. Daily turnover of the blood cells is enormous. Each day, more than 300 billion cells (Notta et al., 2016) composed of $2x10^{11}$ erythrocytes (Dzierzak and Philipsen, 2013), 10^{11} platelets (Kaushansky, 2005) and 10^{11} leukocytes (Wirths et al., 2013) are produced in the human bone marrow by hematopoietic stem cells (HSCs), and multipotent, lineage-restricted progenitors collectively called hematopoietic stem and progenitor cells (HSPCs).

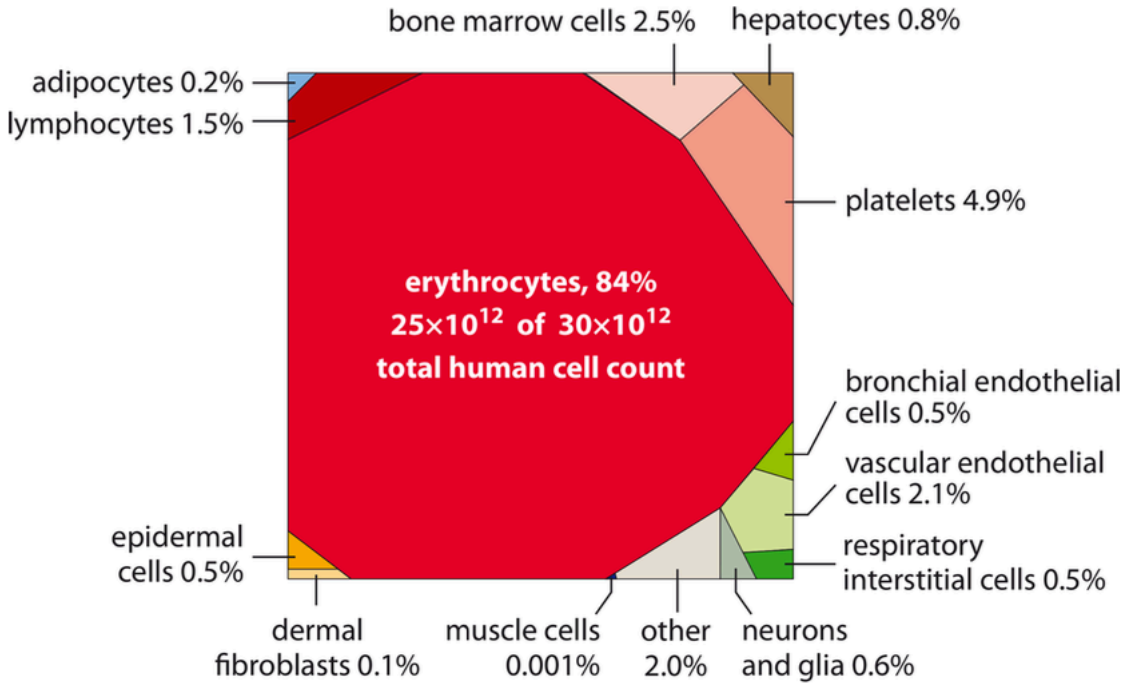


Figure 1. The distribution of the number of human cells by cell type depicted as a Voronoi tree map where polygon area is proportional to the number of cells. Adapted from (Sender et al., 2016).

HSCs are the most-studied adult stem cells owing to the properties that allow their examination in great detail. First of all, these cells are available from the blood; hence their isolation is minimally invasive. Secondly, they are naturally capable of extravasating into the tissues and thus withstand immense shear forces. This makes the isolation and characterization by flow cytometry available. Lastly, HSCs grow single-cell colonies under the proper culture conditions, which allows performing

clonogenicity assays. The above-described properties made detailed characterization of distinct hematopoietic stem cell and progenitor populations possible, by combining the surface marker expression patterns with functional tests examining self-renewal properties (Akashi et al., 2000; Kiel et al., 2005; Kondo et al., 1997; Morrison and Weissman, 1994; Osawa et al., 1996). The exceptional capability of HSCs to repopulate mature blood cell lineages was demonstrated by the experiment, in which a single stem cell was able to restore the entire lymphohematopoietic system of the lethally irradiated mouse (Osawa et al., 1996).
Since 1957, hematopoietic stem cell transplantation is the prime example of success in the field of regenerative medicine (Thomas et al., 1957). It has become a life-saving and routine treatment for patients suffering from blood disorders or hematological malignancies. HSC transplantation reaches up to 22,000 transplants in the US and 40,000 in Europe annually (Norkin and Wingard, 2017; Passweg et al., 2016).

1.1.1 Embryonic hematopoiesis

Development of embryonic hematopoietic system starts once the embryo reaches a size, which makes the delivery of oxygen and other factors by the simple diffusion impossible. Hematopoiesis first takes place the extraembryonic tissues – yolk sac, allantois, and placenta. In mice, embryonic hematopoiesis moves from extraembryonic tissues into the aorta-gonad-mesonephros (AGM) region. AGM and placenta are also the regions, where first cells with the properties similar to the adult HSCs are produced (Gekas et al., 2005; Ivanovs et al., 2011; Medvinsky and Dzierzak, 1996; Müller et al., 1994; Ottersbach and Dzierzak, 2005). Source of HSCs in the mouse embryo is a specialized subset of vascular endothelium that acquires hematopoietic potential, called hemogenic endothelium (Eilken et al., 2009). This observation could be later also confirmed in the developing zebrafish embryos (Bertrand et al., 2010; Kissa and Herbomel, 2010).
Embryonic HSCs also have different properties in comparison to their adult counterparts – they are actively cycling and regenerate hematopoiesis faster and more robustly in order to establish a functional hematopoietic system in a developing embryo. Around day E11, murine HSCs migrate from the hemogenic endothelium to the fetal liver, the major site of hematopoiesis during embryonic development. Later on, at day E16.5 they migrate to the bone marrow (Coşkun et al., 2014). Bone

marrow and to the lesser extent also spleen, become the primary sites of hematopoiesis in the adult organism. First, bone marrow of long bones, such as tibia and femur, is in charge of generating blood cells. Later, the production gradually shifts to vertebrae, sternum, ribs, and cranium (Figure 2).

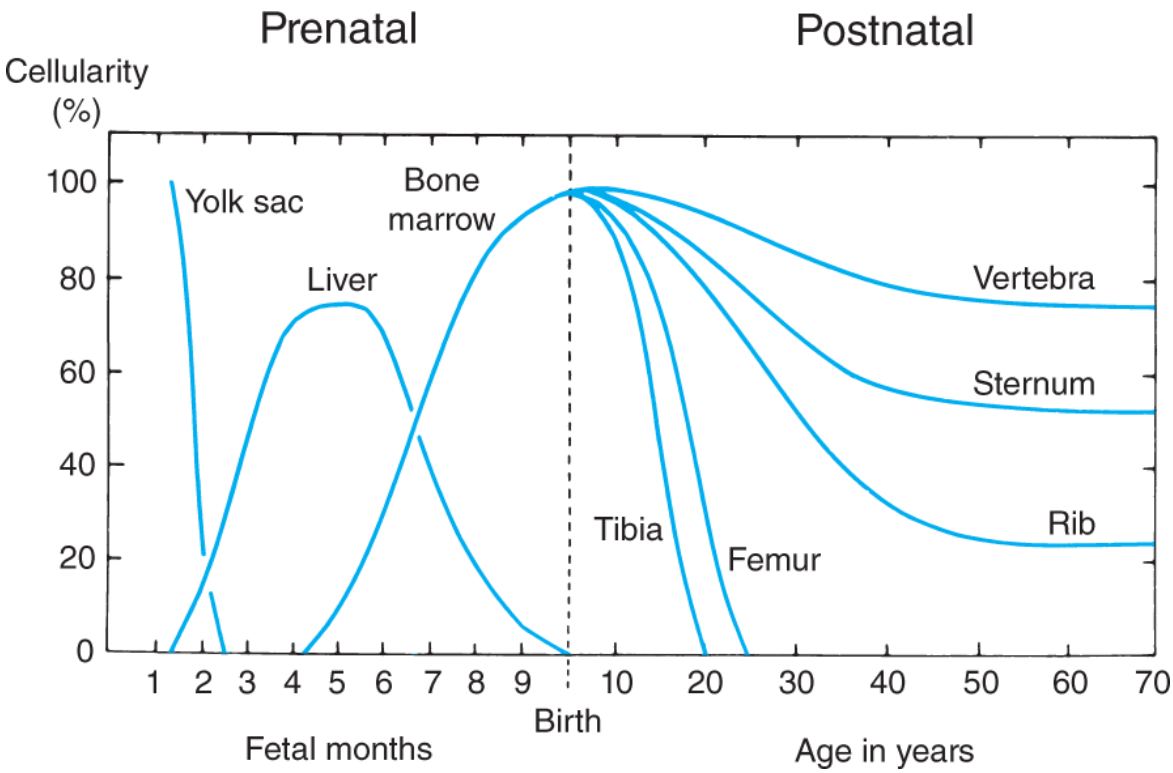


Figure 2. Hematopoietic development in the embryonic, fetal and postnatal stage of development. Blue curves display the relative contribution of the different anatomical sites to the hematopoiesis during ontogeny. Adapted from (Kaushansky et al., 2016).

1.1.2 Adult hematopoiesis

The hematopoietic system is hierarchically organized. At the top of the hierarchy sit HSCs and number of progenitor stages with increasingly restricted lineage potential down to the mature blood cells are located below (Figure 3). Long-term HSCs (LT-HSCs), robustly self-renew and are able to reconstitute all blood lineages upon transplantation into lethally irradiated recipients, thus fulfilling the formal criteria of the stem cell definition: self-renewal, and differentiation potential (Benveniste et al., 2010; Osawa et al., 1996). Functional LT-HSCs in murine models are defined as having the ability to reconstitute all blood cell lineages for more than 24 weeks after the transplantation and, to repopulate the secondary recipients (Benveniste et al., 2010). Many studies used various surface markers to characterize different cell

populations and identified, that stem cell activity in mouse lies within so-called LSK (Lin$^-$/Sca-1$^+$/c-kit$^+$) population (Figure 3).

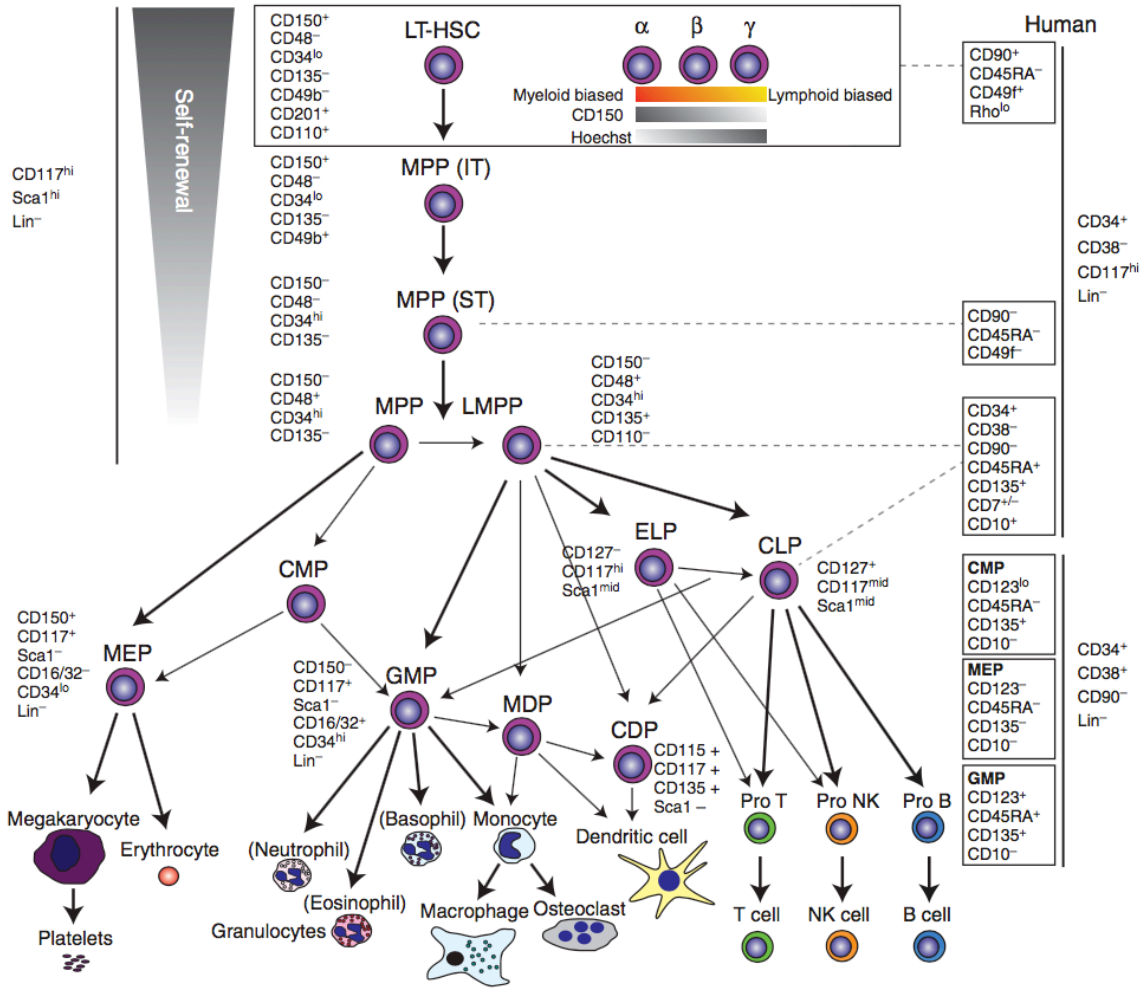


Figure 3. Hierarchy of the hematopoietic system. Long-term hematopoietic stem cells (LT-HSC) that are characterized by the self-renewal and differentiation potential sit at the apex of the hierarchy. Surface markers for prospective isolation of different populations are shown for both mouse and human system (boxed combinations on the right). Bold arrows display established, and thin arrows display alternative differentiation routes. Abbreviations used: HSC, hematopoietic stem cell; MPP, multipotent progenitor; LT-, long-term; IT-, intermediate-term ; ST-, short-term; LMPP, lymphoid-primed MPP; ELP, early lymphoid progenitor; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte–macrophage progenitor; MEP, megakaryocyte–erythrocyte progenitor; CDP, common dendritic progenitor; MDP, monocyte–dendritic cell progenitor; NK, natural killer cell. Adapted from (Rieger and Schroeder, 2012).

LSK population represents ~ 0.06% of the total bone marrow cells and contains cells with the expression of immature multipotent cell markers CD117 (c-kit) and Sca-1, and lack of the expression for the markers of mature lineages (Lin) (Morrison and Weissman, 1994; Okada et al., 1992; Spangrude et al., 1988; Uchida, 1992) LT-HSC population is very rare, and approximately one in thirty LSK cells is LT-HSC (0.002% in BM). The most widely used approach to obtain murine HSCs relies on enriching for LSK/CD150$^+$/CD48$^-$/CD34^{low} population, in which at least 50% of cells are true LT-HSCs with long-term repopulating potential (Kiel et al., 2005). An

alternative strategy to cell surface markers is the use of functional properties of HSCs to isolate them. HSCs are equipped with very active ATP-binding cassette (ABC) transporters, a type of an efflux system used to protect themselves from toxic molecules. HSCs reside in the so-called side population (SP) that effectively transport cell-penetrating dyes such as Hoechst 333420 (Goodell, 1996), or Rhodamine123 dye (Bertoncello et al., 1985) out of the cell.

HSCs need proper microenvironment in order to maintain their functionality. In the bone marrow, they reside in the perivascular niche. Perivascular niche is localized in the proximity of sinusoidal blood vessels that provide support for hematopoietic stem and progenitor cells (Crane et al., 2017; Kunisaki et al., 2013; Nombela-Arrieta et al., 2013). It also controls quiescence, proliferation, and differentiation of HSCs. Under steady-state conditions, ~ two-thirds of HSCs are kept in the quiescent state (Passegué et al., 2005).

HSCs differentiate into a cascade of progenitor cell types with declining multilineage potential before unilineage commitment occurs (Figure 3) (Adolfsson et al., 2005; Kondo et al., 1997; Wilson et al., 2008). Unilineage commitment is the point, at which the cell becomes irreversibly restricted to a defined fate and loses the potential to differentiate. Multiple studies demonstrated the heterogeneity of HSPC compartment and challenged the established hierarchical hematopoietic lineage tree (Figure 4A). For example, a study by Adolfsson and colleagues identified, that LSK/CD34^{+}/Flt3^{+} population failed to produce erythroid and megakaryocytic colonies, and termed it lymphomyeloid-restricted progenitors (LMPP) (Adolfsson et al., 2005) (Figure 4B).

In vivo lineage tracing study showed that common myeloid progenitor (CMP) population is highly heterogeneous, and the divergence between myeloid cells and erythrocytes starts already within the multipotent progenitor pool (MPP) (Perié et al., 2015). Another study, employing single-cell transplantation experiments suggested the so-called myeloid bypass model. In this model, a myeloid-restricted population (MyRP) produced by HSCs can clonally expand and is responsible for the generation of myeloid cells (Figure 4C) (Yamamoto et al., 2013).

Several studies advocate the existence of megakaryocyte-primed HSCs, that are associated with age and mainly responsible for the generation of megakaryocytes and platelets, especially under the stress conditions such as inflammation (Grover et al., 2016; Haas et al., 2015; Sanjuan-Pla et al., 2013).

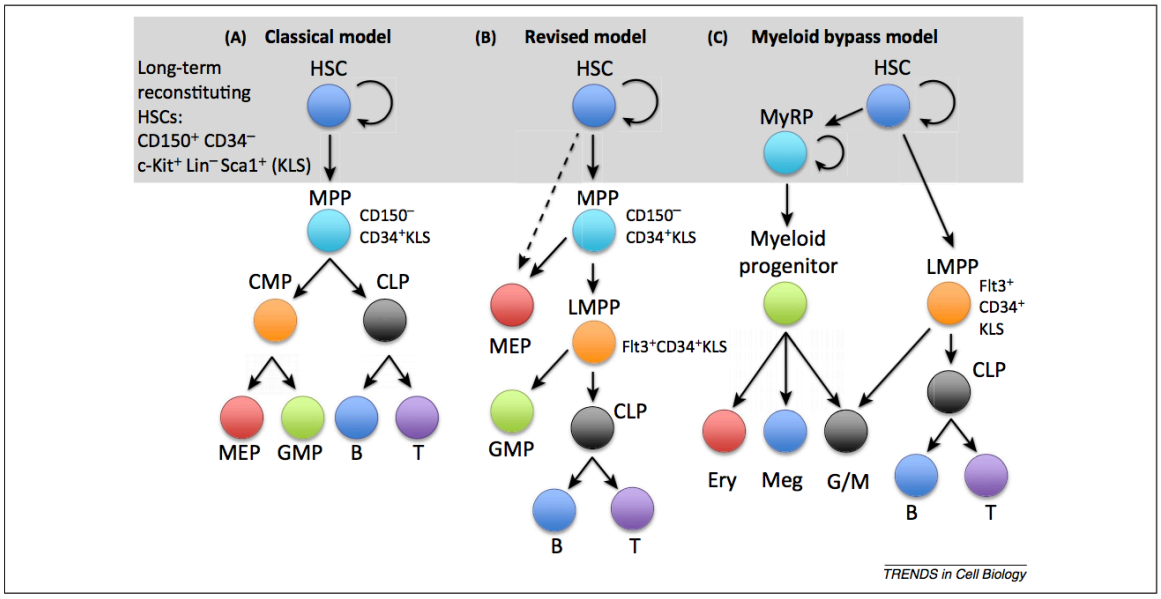


Figure 4. Different models for the lineage commitment. A) The classical model is defined by the cascade in which downstream progenitors are more restricted in their differentiation potential. First branching occurs at the level of common myeloid progenitors (CMP) and common lymphoid progenitors (CLP). B) A revised model incorporating the lymphomyeloid-restricted progenitors (LMPPs) population that has lost the potential to produce megakaryocytes and erythroid progenitors (Adolfsson et al., 2005). C) Myeloid bypass model suggests the existence of myeloid-restricted progenitors (MyRPs) that are directly derived from LT-HSCs and retain self-renewal capacity. This model proposed that self-renewal and differentiation capacity can be uncoupled (Yamamoto et al., 2013). Abbreviations: Ery, erythroid; G/M, granulocyte/monocyte; GMP, granulocyte/monocyte progenitor; HSC, hematopoietic stem cell; Meg, megakaryocyte; MEP, megakaryocyte/erythroid progenitor; MPP, multipotent progenitor. Adapted from (Nimmo et al., 2015).

All these studies have used lethal irradiation and in their experimental setup. Irradiation is known to create a cytokine release syndrome, that favors engraftment. Thus, these models represent stress-induced hematopoiesis. To overcome this limitation, several groups developed alternative approaches to monitor unperturbed, steady-state hematopoiesis (Busch et al., 2015; Rodriguez-Fraticelli et al., 2018; Sun et al., 2014). These studies showed that steady-state hematopoiesis is driven by a large number of long-lived progenitors rather than the small number of multipotent hematopoietic cells (Busch et al., 2015; Sun et al., 2014), and that megakaryocyte lineage arises largely independently of other hematopoietic fates (Rodriguez-Fraticelli et al., 2018).

Human HSCs express CD34 molecules on their surface and together with the lack of CD38 and lineage markers define human HSPC population: $CD34^+/CD38^-/Lin^-$. However, due to the technical limitations, purification of human HSPC compartment is far less advanced and thus $CD34^+/CD38^-/Lin^-$ population contains <1% of functional HSCs. By adding further specific surface markers, scientists were able to further enrich the HSPC population to about 15% (Notta et al., 2011). It is important to mention that the abundance of human HSCs in the purified fraction may be

underestimated since murine microenvironment is less permissive for engraftment of human HSCs, and therefore new xenograft models are being developed (Rongvaux et al., 2014). Human hematopoietic differentiation hierarchy was thought to be similar to the murine one (Majeti et al., 2007). However, a recent study challenged the existence of oligopotent progenitors using single-cell assays (Notta et al., 2016).

1.2 Hematopoietic growth factors and their receptors

Hematopoiesis is regulated by the hematopoietic growth factors that are required for the survival and proliferation of blood cells at all stages of the development. Growth factors bind to the extracellular domains of cognate receptors, form signaling complexes with receptors and induce intracellular signaling events leading to the changes in gene expression. Cytokine receptor superfamily is a large group of transmembrane receptors that share a typical structure of around 200 amino acids in their extracellular domain. This superfamily has been subdivided into 5 families based on the characteristic structural motif: type I cytokine receptor family (also called hematopoietin receptor family), type II (also known as interferon receptor family), tumor necrotic factor (TNF) receptor family, immunoglobulin receptor family, and chemokine receptor family. The majority of growth factors that modulate the activity of the hematopoietic system bind to the type I cytokine receptor family (Figure 5). Lineage-specific growth factors act on a particular blood lineage due to the restricted expression of cytokine receptors and include erythropoietin (EPO), thrombopoietin (TPO), and granulocyte colony-stimulating factor (G-CSF). Multilineage hematopoietic growth factors include granulocyte-macrophage colony-stimulating factor (GM-CSF), and interleukins. Both type I and type II receptors lack catalytic activity and thus require Janus kinase (JAK) proteins that possess phosphorylation activity (Figure 5). Erythropoietin and thrombopoietin will be discussed further in more detail in the following chapters.

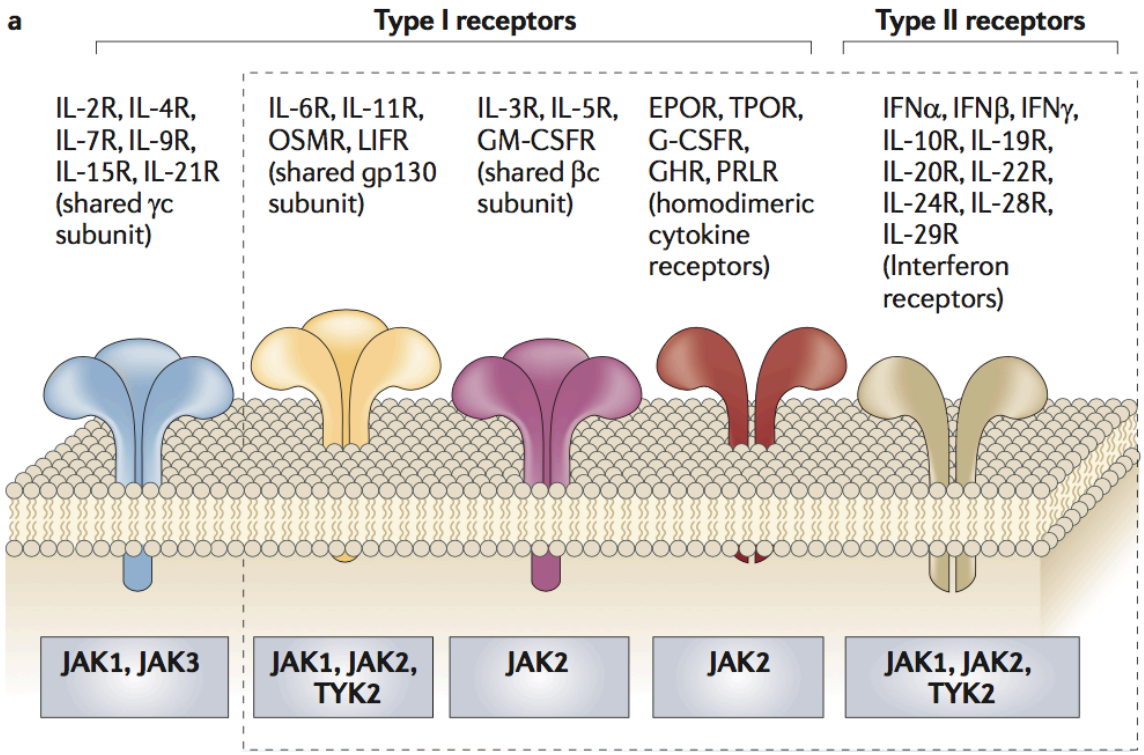


Figure 5. Cytokine receptors and JAK kinase family. Type I and type II receptors are classified according to the three-dimensional structure of their ligands. These receptors lack intrinsic kinase activity and depend on receptor-associated Janus kinases (JAKs) to transmit the signal. Type I cytokine receptor family (also called hematopoietin receptor family) binds growth factors that are modulating the activity of the hematopoietic system. Type II receptors are mostly bound by interferons and involved in the regulation of the immune response. JAK kinase family consists of 4 different kinases (JAK1, JAK2, JAK3, and TYK2) that are differentially activated by different cytokines. Adapted from (Quintás-Cardama et al., 2011).

1.2.1 Hematopoietic growth factor-induced signaling

Hematopoietic cytokines bind to their respective receptors to initiate the cascade of signaling events leading to the changes in the expression of target genes. The central axis for this signaling is the JAK-STAT pathway, that was discovered by the studies on gene induction by interferon (Darnell et al., 1994). Janus kinase family consists of four members: JAK1, JAK2, JAK3, and a non-receptor tyrosine kinase TYK2, that are differentially activated by different cytokines (Figure 5) (Stark et al., 1998). All JAK family members have a common domain structure design containing four-point-one, ezrin, radixin, moesin (FERM) domain at its N-terminus that connects JAKs with the receptors, Src homology 2 (SH2) domain, and JAK-homology domains 1 (JH1) and 2 (JH2) (Pearson et al., 2000). JH1 domain locates to the C-terminus of the protein and possesses the kinase activity (Shuai et al., 1993). JH2 is also termed “pseudokinase” domain, as it was originally thought not to have any enzymatic activity. However, a study by Ungureanu and others

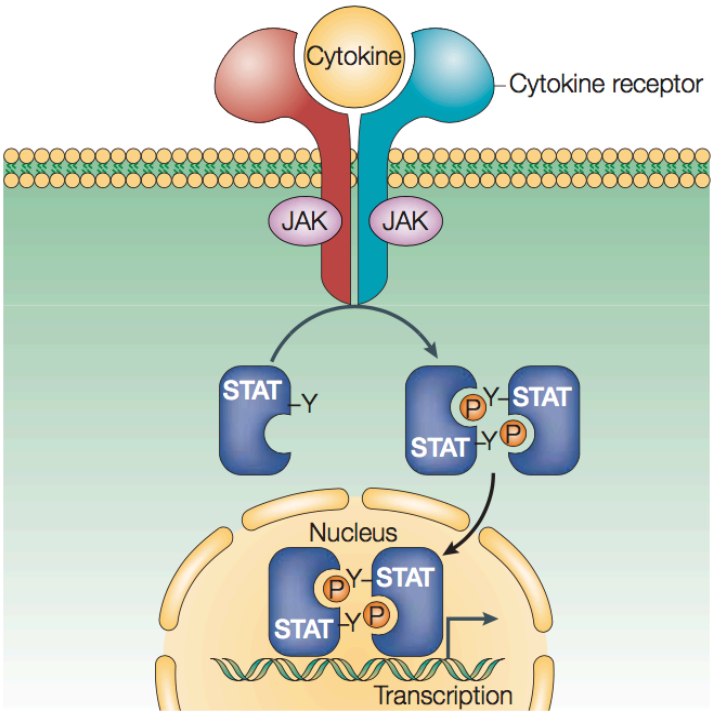


Figure 6. JAK-STAT signaling pathway. The activation of Janus kinase (JAK) upon binding of the cytokine to the receptor results in the phosphorylation of signal transducer and activator of transcription (STAT) proteins that dimerize and translocate to the nucleus to take part in the gene transcription activation. Adapted from (Shuai and Liu, 2003).

demonstrated that JH2 domain of JAK2 has a weak dual- type kinase activity (Ungureanu et al., 2011). Ligand binding causes the receptor dimerization and the activation of JAK kinases that are constitutively associated with the intracellular domains of the receptor. JAK kinase undergoes a conformational change leading to the auto-phosphorylation of specific tyrosine residues. This activation promotes the recruitment of signal transducer and activator of transcription (STAT) transcription factors that are also phosphorylated by JAK kinases. STAT protein family consists of seven members: STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B, and STAT6 that share a high degree of homology across several conserved domains, including SRC homology 2 (SH2) domain mediating activation and dimerization of STATs (Fu and Zhang, 1993; Shuai, 1994), DNA-binding domain (Horvath et al., 1995), and a transactivation domain (Shuai et al., 1993). STAT phosphorylation causes dimerization through their SH2 domains and subsequent translocation to the nucleus, where they bind to the specific promoter DNA sequences and activate the transcription of target genes (Figure 6) (Darnell, 1997; Levy and Darnell, 2002). In addition, JAK kinases also activate MAPK and PI3K/Akt signaling axes.

The JAK-STAT signaling pathway is negatively regulated predominantly by suppressor of cytokine signaling (SOCS) family of proteins, consisting of eight members: SOCS1-SOCS7 and cytokine-inducible SH2-containing protein (CISH) (Alexander, 2002; Hilton, 2001). SOCS proteins inhibit JAK-STAT signaling through distinct mechanisms. SOCS1 binds directly to JAKs via its SH2 domain, hence directly inhibiting the JAK activity (Endo et al., 1997; Naka et al., 1997; Starr et al.,

1997). SOCS3 needs to bind to the activated receptor in order to exert its inhibitory function (Sasaki et al., 2000). Finally, CISH is competing with STATs for the receptor binding sites (Yoshimura, 1998).

1.2.2 JAK-STAT signaling in hematopoiesis

The essential role of JAK2 in hematopoiesis is demonstrated by the embryonic lethality of homozygous germline deletion of JAK2, due to the defective erythropoiesis (Neubauer et al., 1998; Parganas et al., 1998). JAK2 associates with multiple cytokine receptors: erythropoietin receptor (EPOR), thrombopoietin receptor (MPL), granulocyte-colony stimulating factor receptor (G-CSFR), granulocyte-macrophage colony-stimulating factor receptor (GM-CSFR) and interleukin-3 receptor (IL-3R) being the most prominent. Failure of liver progenitors from JAK2-deficient mice to respond to the stimulation with EPO, TPO, GM-CSF or IL-3 highlights the importance of JAK2 in mediating the cytokine-induced signaling (Parganas et al., 1998).

JAK2 activates STAT1, STAT3, STAT5a, and STAT5b. STAT1-deficient mice do not have any developmental abnormalities, but possess an attenuated response to viral infections due to the disrupted interferon signaling (Durbin et al., 1996). These mice also show delayed erythroid differentiation, decrease in erythroid colony-forming units (CFU-E), and defective polyploidization of megakaryocytes (Halupa et al., 2005; Huang et al., 2007). Constitutional STAT3 deletion caused defects in visceral endoderm and resulted in the embryonic lethality (Takeda et al., 1997). Further studies of conditional STAT3 deletion found its role in T-cell proliferation and innate immunity regulation (Takeda et al., 1998; Welte et al., 2003). STAT5a and STAT5b are closely related and partially redundant proteins essential for hematopoiesis. STAT5a/STAT5b double knock-out embryos die from severe anemia (Teglund et al., 1998), revealing the essential role of STAT5 isoforms in the EPOR-mediated antiapoptotic signaling (Socolovsky et al., 1999).

Erythropoietin signaling depends on the JAK2-STAT5 axis since mice lacking the expression of EPO receptor (Wu et al., 1995), JAK2 (Neubauer et al., 1998; Parganas et al., 1998), or STAT5 (Teglund et al., 1998) show inefficient erythropoiesis and embryonic lethality. Both JAK2-STAT3 and JAK2-STAT5 pathways mediate thrombopoietin signaling (Miyakawa et al., 1996). Granulocytic

differentiation via G-CSF/G-CSFR axis is mostly transmitted through JAK1 and only to a lesser extent through JAK2 (Shimoda et al., 1997).

In 2005, four groups independently reported the discovery of a gain-of-function mutation in JAK2 in patients with myeloproliferative neoplasms (Baxter et al., 2005; James et al., 2005; Kralovics et al., 2005; Levine et al., 2005) The mutation is a single nucleotide G→T transversion that results in a change from valine to phenylalanine at the position 617 (V617F). This substitution disrupts the inhibitory activity of JH2 pseudokinase domain, that has significant homology to the tyrosine kinase domain (JH1) and possess a weak dual-type kinase activity (Ungureanu et al., 2011). As a consequence, JH1 domain is constantly activated and hematopoietic cells do not depend anymore on the cytokine signaling to instruct their proliferation. Noteworthy, the disruption of ATP binding site in the pseudokinase domain JH2 of wildtype JAK2 has a minimal effect. However, it can suppress the hyperactivation of mutated JAK2 (Hammaren et al., 2015).

1.3 Erythropoiesis

Red blood cells perform an essential role in maintaining the existence of all vertebrate organisms, owing to their oxygen delivery function. Erythrocyte-committed progenitors undergo an enormous expansion to fulfill the demand for $2x10^{11}$ new erythrocytes each day (Dzierzak and Philipsen, 2013). In the steady-state hematopoiesis, around 1% of erythrocytes are replaced every day. This process is governed by erythropoietin, a hormone produced in the kidneys, whose production is tightly regulated. Developmental studies identified two distinct types of erythrocytes – embryonic and adult ones. They develop at different anatomical sites from distinct types of hematopoietic progenitors.

1.3.1 Erythropoiesis during fetal and embryonic development and postnatal life

Erythrocytes are formed in two waves, a primitive wave followed by the definitive wave. Primitive erythropoiesis occurs in the yolk sac. Erythroid cells are derived from mesodermal cells that migrate to the yolk sac and come in close contact with endodermal cells (Kinder et al., 1999; Lawson et al., 1991). In the yolk sac, they form blood islands that contain large and enucleated primitive erythrocytes around day E7.5 in the mouse and between day 16 and day 20 of gestation in humans

(Tavian and Peault, 2005). Primitive erythropoiesis is characterized by the expression of embryonic globins ε, γ, and ζ that form a variety of tetramers. Definitive erythropoiesis occurs first in the fetal liver. This event is marked by the switch to the fetal hemoglobin HbF, that is composed of $\alpha_2\gamma_2$ tetramers. Around the time of the birth, fetal hemoglobin expression is silenced, and erythropoiesis moves to the bone marrow and spleen. Adult hemoglobin composed of $\alpha_2\beta_2$ (HbA1) or $\alpha_2\delta_2$ (HbA2) tetramers becomes responsible for oxygen transport (Sankaran et al., 2010). Erythroid cells are derived from HSCs and follow a differentiation path towards mature erythrocytes through multiple stages of progenitors (Figure 3).

Most immature erythroid progenitors are called burst forming units-erythroid (BFU-E) and give rise to a multi-subunit colony containing up to a few thousands of hemoglobinized cells in the methylcellulose culture (Gregory and Eaves, 1977). Their growth is dependent on stem cell factor (SCF), IL-3, and EPO. BFU-E progenitors appear in the bone marrow and to the lesser extent also in the peripheral blood (Figure 7). Following stage are colony forming units-erythroid (CFU-E). They form small colonies consisting of 10-120 cells appearing after 2-3 days (mouse) and 5-8 days (human) in the methylcellulose. They are dependent entirely on EPO as a growth factor and do not appear in the peripheral blood under the normal circumstances (Koury and Bondurant, 1990a).

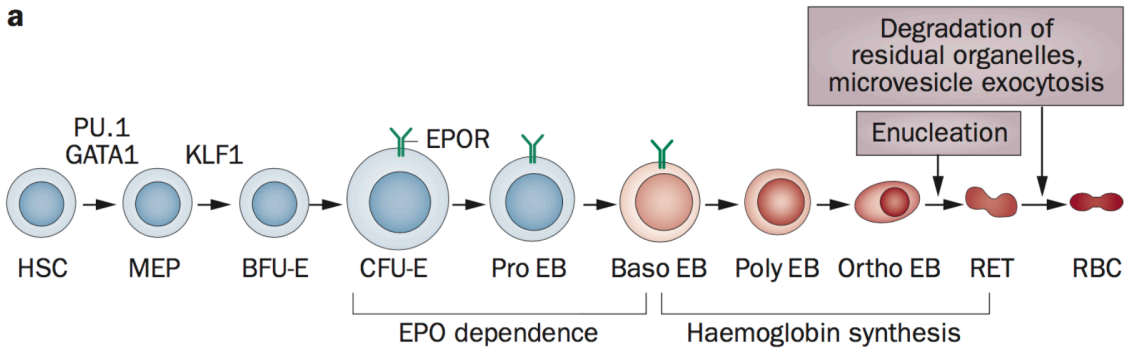


Figure 7. Scheme of erythroid differentiation illustrates progressive stages of erythroid progenitors towards the mature erythrocytes. Relative cell sizes and presumed or known morphologic appearances at various stages are shown. PU.1 and GATA are transcription factors determine the commitment of HSC towards the erythroid fate. Transcription factor KLF-1 determines whether MEP will differentiate towards the erythroid or megakaryocyte progenitor. CFU-E progenitors and proerythroblasts (Pro EB) are dependent on the EPO signaling. Stages of hemoglobin synthesis display relative accumulations of hemoglobin as increasing intensity of red in the cytoplasm. Abbreviations: HSC, hematopoietic stem cell; MEP, megakaryocyte-erythroid progenitor; BFU-E, burst-forming unit, erythroid; CFU-E, colony-forming unit, erythroid; Pro EB, proerythroblast; Baso EB, basophilic erythroblast; Poly EB, polychromatophilic erythroblast; Ortho EB, orthochromatic erythroblast; RET, reticulocyte; RBC, red blood cell; EPOR, erythropoietin receptor. Adapted from (Koury and Haase, 2015).

CFU-E progenitors undergo a series of rapid cell divisions through different maturation steps: proerythroblast, basophilic erythroblast, polychromatophilic erythroblast and orthochromatic erythroblast (Dzierzak and Philipsen, 2013). This progressive erythroid precursor maturation is characterized by the decrease in cell size and accumulation of hemoglobin. Hemoglobin synthesis begins in the late basophilic erythroblast and continues through to the reticulocyte stage. At the terminal stage of erythrocyte differentiation, cells lose their ability to proliferate, perform nuclear condensation and expel the nucleus before they enter the circulation as reticulocytes. Reticulocyte maturation is a complex process characterized by 20% loss of plasma membrane surface area, reduction of cell volume and loss of cytoplasmic organelles (Johnstone, 1992; Ney, 2011). It takes reticulocytes around one week to fully mature into the bi-concave shaped erythrocytes with a lifespan of around 115 days (Franco, 2012). Senescent erythrocytes are phagocytosed by splenic, hepatic or bone marrow macrophages (Willekens et al., 2008). Erythroid differentiation takes place in the bone marrow in the structural unit called erythroblastic island (Chasis and Mohandas, 2008). Erythroblastic island is composed of a central macrophage, also called a nurse cell and erythroid cells at various stage of differentiation (Figure 8).

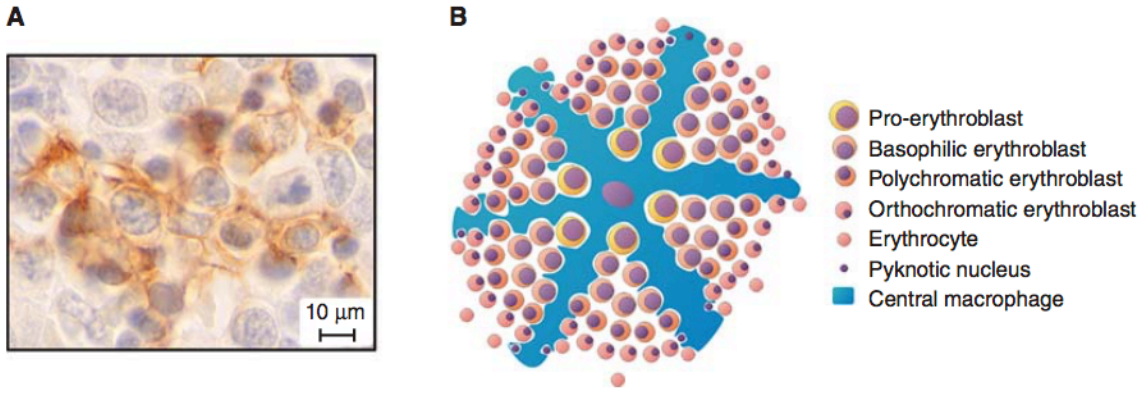


Figure 8. An erythroblastic island. A) Immunohistology image of E13.5 mouse fetal liver shows a central macrophage stained with the F4/80 antibody in brown surrounded by erythroid cells at various stages of differentiation. B) Schematic drawing of an erythroblastic island. Reticulocytes detach from the macrophage before leaving the bone marrow and entering the circulation. Adapted from (Dzierzak and Philipsen, 2013).

Cell death signaling also plays an important role in erythroid differentiation. Up to 60% of erythroblasts commit apoptosis in the mouse spleen under steady-state conditions (Liu et al., 2006). Signaling induced by erythropoietin prevents apoptosis of CFU-Es, proerythroblasts, and early basophilic erythroblasts (Koury and Bondurant, 1990b; Muta and Krantz, 1993; Wu et al., 1995). Individual erythroid progenitors have a significantly different sensitivity to erythropoietin that is

associated with varying expression of the CD95 (FAS) death receptor during the differentiation process. Upon binding of CD95L (FASL), a ligand of CD95 receptor, produced by the mature erythroblasts within the erythroblastic island (De Maria et al., 1999), progenitor cells activate caspase-3 dependent pro-apoptotic signaling pathways (Gregoli and Bondurant, 1999; Rubiolo et al., 2006). This negative feedback loop efficiently controls the rate of erythrocyte production (Kelley et al., 1993).

1.3.2 Transcriptional regulation of erythropoiesis

Multiple transcription factors are involved in the establishment of the erythroid lineage. GATA1 plays an essential role in the expression of erythroid-specific genes, as GATA1 deficiency causes the arrest at the proerythroblast stage (Fujiwara et al., 1996; Pevny et al., 1995). Different functional domains of GATA1 are required for the activation of target genes in primitive and definitive stages of erythropoiesis, suggesting GATA1 is a part of different transcriptional complexes (Shimizu et al., 2001).

Another important transcription factor is KLF1, that is responsible for the terminal erythroid differentiation (Miller and Bieker, 1993). KLF1 expression is largely restricted to the erythroid cell lineage (Southwood et al., 1996). Homozygous deletion of KLF1 in mice causes fatal anemia in the developing fetus around day E14, due to the lack of β-globin expression, when embryos switch to the definitive erythropoiesis in the fetal liver (Nuez et al., 1995; Parkins et al., 1995). Importantly, the number of CFU-E cells in the KLF1 homozygous mutants is similar to their wildtype siblings, indicating that the KLF1 deficiency causes a very late erythroid defect (Nuez et al., 1995; Parkins et al., 1995). Moreover, KLF1-deficient erythroid progenitors have abnormal morphology and fail to expel the nuclei, indicating that β-globin is not the only target of KLF1. Indeed, genome-wide expression analysis of KLF1-deficient erythroid cells showed that KLF1 activates multiple erythroid-specific genes including fetal and adult globins (Basu et al., 2007; Hodge et al., 2006; Merryweather-Clarke et al., 2011).

1.3.3 Erythropoietin (EPO) and erythropoietin receptor (EPOR)

A humoral “factor” that regulates erythropoiesis was first described by Carnot and Deflandre as a result of an experiment, in which increased red blood cell counts

were observed in normal rabbits after the infusion with the serum of anemic animals (Carnot and Deflandre, 1906). Later confirmed by Erslev (Erslev, 1953), these studies led to the hypothesis about circulating erythroid-stimulating hormone, termed erythropoietin (EPO). Organ ablation studies in rats (Jacobson et al., 1957) and man (Mirand et al., 1969; Nathan et al., 1964) further established kidney as the major site of EPO production. This led to the isolation of EPO protein from the urine of aplastic anemia patients (Miyake et al., 1977) and to the molecular cloning of the *EPO* gene (Jacobs et al., 1985; Lin et al., 1985).

2901 base pairs long human *EPO* gene localizes to the chromosome 7q22.1, contains 5 exons and produces a 193-residue polypeptide with the molecular weight of 30.4 kDa. Heavy glycosylation that accounts for 40% of the molecular weight, slightly impedes its biological activity. However, it is critical for maintaining long half-life in the serum, that reaches around 7-8 hours (Goldwasser et al., 1974). EPO binds with a high affinity (~100pM) to the cognate EPO receptor (EPOR) that is present in relatively small numbers (~1000 molecules/cell) on the surface of erythroid progenitors (Bunn, 2013).

Human erythropoietin receptor encoded by 6 exons on chromosome 19p13.2 produces a 508-amino acid polypeptide with a molecular weight of 55 kDa It is a member of the type I cytokine receptor family (Figure 5) and is present on the cell membrane as a homodimer, even in the absence of the ligand (Livnah et al., 1999). Binding of EPO causes the conformational change that brings its intracellular domains into the proximity (Figure 9). This leads to the initiation of JAK2-mediated intracellular signaling events (described in detail in chapter 1.2.1) and activation of several signaling axes JAK/STAT, PI3K/Akt, ERK/MAPK and PKC pathways (Richmond et al., 2005; Witthuhn et al., 1993; Youssoufian et al., 1993). Erythropoietin-EPOR complexes are subsequently degraded by endocytosis (Gross and Lodish, 2006).

During embryonic development, erythropoietin production takes place in the liver, in hepatocytes surrounding the central vein (Koury et al., 1991) and Ito presinusoidal cells are (Maxwell et al., 1994). Around the birth, EPO synthesis gradually moves to the kidneys that become the primary site of the EPO in adult life. Although this switch is well documented in multiple species (Bondurant et al., 1991; Dame et al., 1998; Eckardt et al., 1992; Peschle et al., 1975; Zanjani et al., 1981), molecular mechanisms leading to this switch remain elusive.

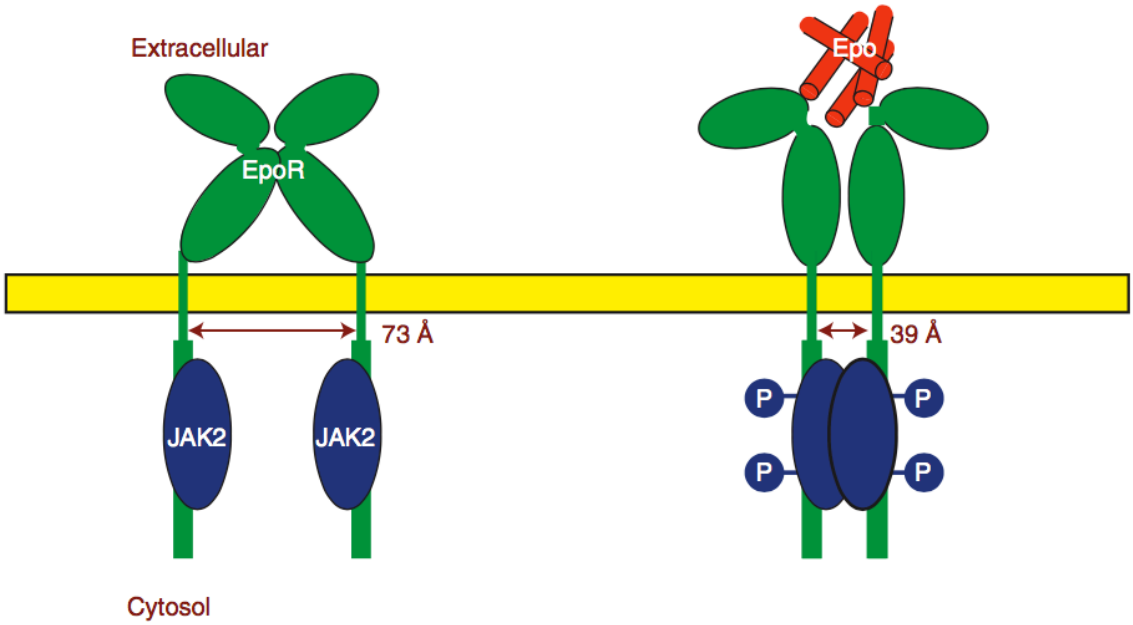


Figure 9. EPO-mediated signaling. EPO receptor (EPOR) is present on the cell surface even in the absence of the ligand. Binding of EPO causes a conformational change that brings intracellular domains of the receptor into close proximity. This causes the transphosphorylation of associated JAK2 kinase and initiates the signal transduction cascade. Adapted from (Bunn, 2013).

It has been suggested that the spatio-temporal expression of GATA transcription factors may play a role in this process (Weidemann and Johnson, 2009). GATA proteins belong to the family of zinc-finger DNA-binding proteins and play critical roles in cell growth and differentiation (Lentjes et al., 2016). The liver does not contribute to the plasma erythropoietin pool under physiological conditions in adults. In severe hypoxia, however, hepatocytes are capable of synthesizing EPO that accounts for the majority of systemic erythropoietin of non-renal origin (Fried, 1972; Fried et al., 1969; Mirand et al., 1969; 1968; Obara et al., 2008).

The constant ratio between blood perfusion and oxygen consumption, together with low tissue pO_2 and stable corticomedullary oxygen gradient, make kidney an ideal organ for sensing pO_2 concentration (Wenger and Hoogewijs, 2010). Although several studies proposed that erythropoietin is produced by renal epithelial cells (Loya et al., 1994; Maxwell et al., 1990; Mujais et al., 1999), genetic models combined with immunohistochemistry and *in situ* hybridization identified peritubular interstitial cells on the medulla-cortex boundary as the physiological source of erythropoietin (Figure 10) (Bachmann et al., 1993; Koury et al., 1988; 1989; Lacombe et al., 1988; Obara et al., 2008). Several studies showed that the number of renal EPO-producing cells (REPs) increases proportionally to the degree of hypoxia, and therefore the number of REPs determines the renal contribution to the plasma EPO levels (Obara et al., 2008; Souma et al., 2013; Yamazaki et al., 2013).

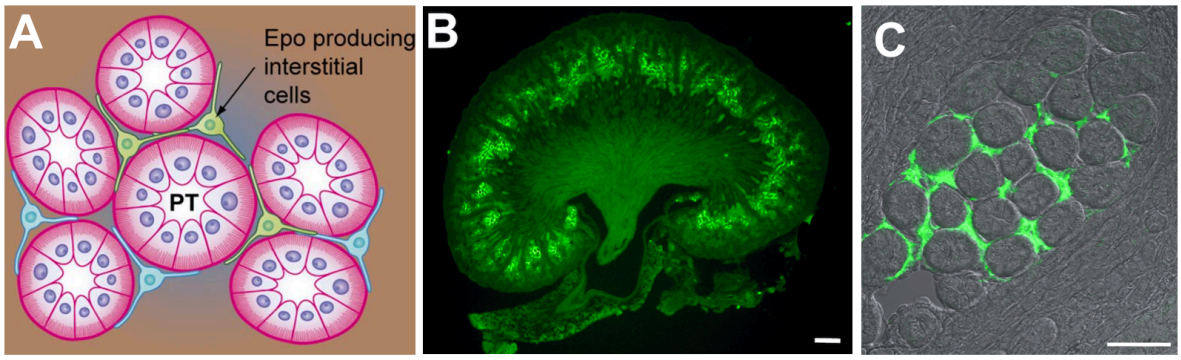


Figure 10. Renal EPO-producing cells (REPs). Panel A illustrates a schematic picture of proximal tubules (PT) formed by epithelial cells. Peritubular interstitial cells that produce EPO are depicted in green. Panel B shows that GFP-labeled REPs are located on the medulla-cortex boundary. Panel C displays a close-up look on the proximal tubule region with peritubular interstitial cells labeled in green. Panel A is adapted from (Noguchi, 2008) and panels B and C from (Pan et al., 2011).

In addition to the kidney and liver, *EPO* expression was also detected in other tissues, such as neurons and glial cells, bone marrow cells, osteoblasts, spleen, hair follicles, lung tissue, cardiomyocytes, and reproductive organs (Bernaudin et al., 2000; Bodo et al., 2007; Dame et al., 1998; Fandrey and Bunn, 1993; Marti et al., 1996; Miro-Murillo et al., 2011; Rankin et al., 2012; Yasuda et al., 1998). These organs are thought not to contribute significantly to the systemic erythropoietin pool, but rather to exert local, non-hematopoietic actions, such as angiogenesis, brain development, and wound healing (Arcasoy, 2008; Jelkmann, 2007).

1.3.4 Erythropoietin regulation

Erythropoiesis is stimulated by hypoxia that activates hypoxia-inducible factors (HIF). HIFs are the master regulators of cellular response to the hypoxic stimulus. In addition to the stimulation of renal end hepatic erythropoietin production, they also promote the uptake and utilization of iron and modify the bone marrow microenvironment to support the maturation and proliferation of erythroblasts. The discovery of human hepatoma cell lines, Hep3B, and HepG2 that up-regulate *EPO* expression in a hypoxia-dependent manner greatly facilitated the initial studies of *EPO* gene regulation (Goldberg et al., 1987).

The *EPO* gene promoter contains a GATA-binding motif with the core sequence 5'-A**GATAAC**A-3', that facilitates the binding of GATA proteins. It was reported that human GATA-1, GATA-2, and GATA-3 are able to bind to the core *EPO* promoter sequence, inhibit the formation of pre-initiation complexes and thereby reduce the expression levels of *EPO* mRNA in Hep3B cells. (Aird et al., 1994; Imagawa et al., 1997; 2002). In addition, GATA-2 and GATA-3 were shown to constitutively repress

EPO in renal tubular cells in the kidney *in vivo* (Obara et al., 2008). Thus, it has been suggested that the GATA box acts as a negative regulatory element that represses the transcription during normoxia. In contrast, GATA-4 was proposed to have a positive effect on *EPO* expression in Hep3B cells, since knockdown of GATA-4 negatively influenced *EPO* mRNA levels (Dame et al., 2004). In hypoxia, GATA binding markedly decreases, which allows for a significant increase in *EPO* gene expression. Since *EPO* promoter has a weak activity, an enhancer element is necessary to upregulate *EPO* expression under hypoxic conditions.

Enhancer elements differ between kidney and liver. In the kidney, an essential regulatory element for renal *EPO* expression is located between -14kb and -6kb upstream of the *EPO* gene (Semenza et al., 1990). A recent study identified HIF binding site 9248 base pairs upstream of *EPO* transcriptional start site and provided *in vitro* evidence that it acts as a HIF-2 dependent distal hypoxia response element (HRE) (Storti et al., 2014). Since 180 kilobase BAC transgene is required to recapitulate endogenous *EPO* expression in the transgenic mouse model (Obara et al., 2008; Yamazaki et al., 2013), there are probably multiple upstream regulatory elements existing.

In the liver, HRE is located in the 3'-region downstream of the *EPO* gene and contains classical hypoxia enhancer that is essential for the hypoxic induction of EPO in the hepatocytes (Blanchard et al., 1992; Semenza et al., 1991; Suzuki et al., 2011). HIF transcription factors bind to the consensus sequence 5'-TACGTGCT-3' within this regulatory element, cooperate with hepatocyte nuclear factor (HNF-4) and the transcriptional co-activator p300 to initiate the transcription of hypoxia-response genes (Semenza et al., 1991).

HIF transcription factors are heterodimeric proteins consisting of an oxygen-sensitive α-subunit and constitutively expressed β-subunit, also called aryl hydrocarbon receptor nuclear translocator (ARNT). To date, 3 distinct α-subunits were identified in man and named HIF-1α (Wang et al., 1995), HIF-2α (Tian et al., 1997), and HIF-3α (Gu et al., 1998). HIF-1α is ubiquitously expressed, whereas HIF-2α has tissue-restricted expression. Most of the studies have focused on HIF-1 and HIF-2 that are mainly involved in the hypoxia-stimulated cellular response. HIF-1α is ubiquitously expressed, whereas HIF-2α has tissue-restricted expression. Although HIF-1 and HIF-2 regulate expression of many genes (Schödel et al., 2011),

HIF-2 is the isoform responsible for the erythropoietin synthesis and iron metabolism (Gruber et al., 2007; Kapitsinou et al., 2010; Mastrogiannaki et al., 2009; Rankin et al., 2007; Scortegagna et al., 2005).

HIF-α subunits are being constitutively synthesized. Therefore, control of the degradation rate is a key step in the erythropoiesis regulation. HIF-α is remarkably unstable under normal oxygen levels (Huang et al., 1996). Oxygen-dependent HIF-α degradation is initiated by prolyl-4-hydroxylase domain (PHD) oxygenases; PHD1, PHD2, and PHD3 (also known as EGLN2, EGLN1, and EGLN3). PHD2 is the main oxygenase responsible for the degradation of HIF proteins under the normoxia. These enzymes use molecular oxygen to hydroxylate HIFα subunits at specific proline residues, Pro402, and Pro564 in human HIF-1α; Pro405, and Pro531 in human HIF-2α (Bruick and McKnight, 2001; Epstein et al., 2001; Hon et al., 2002; Ivan et al., 2001; Jaakkola et al., 2001; Yu et al., 2001). β-domain of von Hippel-Lindau (VHL) protein binds to the proline-hydroxylated HIF-α that is subsequently recognized by VHL-E3-ubiquitin ligase complex. This complex polyubiquitinates HIF-α and targets it for proteasomal degradation. In hypoxic conditions, hydroxylation of HIFα subunits is inhibited, HIF heterodimers translocate to the nucleus and activate the transcription of hypoxia-response genes (Figure 11). Lack of HIF expression leads to anemia (Gruber et al., 2007; Kapitsinou et al., 2010; Scortegagna et al., 2003), whereas impaired regulation of its activity results in erythrocytosis (Percy et al., 2008; Rankin et al., 2007).

In addition to the regulation of hypoxia-stimulated erythropoietin production, HIF-2 also plays a critical role in the regulation of iron uptake. It upregulates transcription of a divalent metal transporter (*DMT1*) and duodenal cytochrome *b* reductase 1 (*DCYTB*) genes (Shah et al., 2009). DCYTB reduces ferric iron (Fe^{3+}) to its ferrous form (Fe^{2+}) and DMT1 transports the iron from the gut lumen to the cytoplasm of the cells. HIF proteins also regulate the expression of transferrin (*TF*), which transports iron in the plasma (Rolfs et al., 1997), transferrin receptor (*TFRC*) involved in the iron uptake by the cells (Lok and Ponka, 1999), heme-oxygenase-1 (*HO-1*) that is responsible for the iron recycling from phagocytosed erythrocytes (Lee et al., 1997), and ferroportin (*FPN*), the only known cellular iron exporter (Taylor et al., 2011).

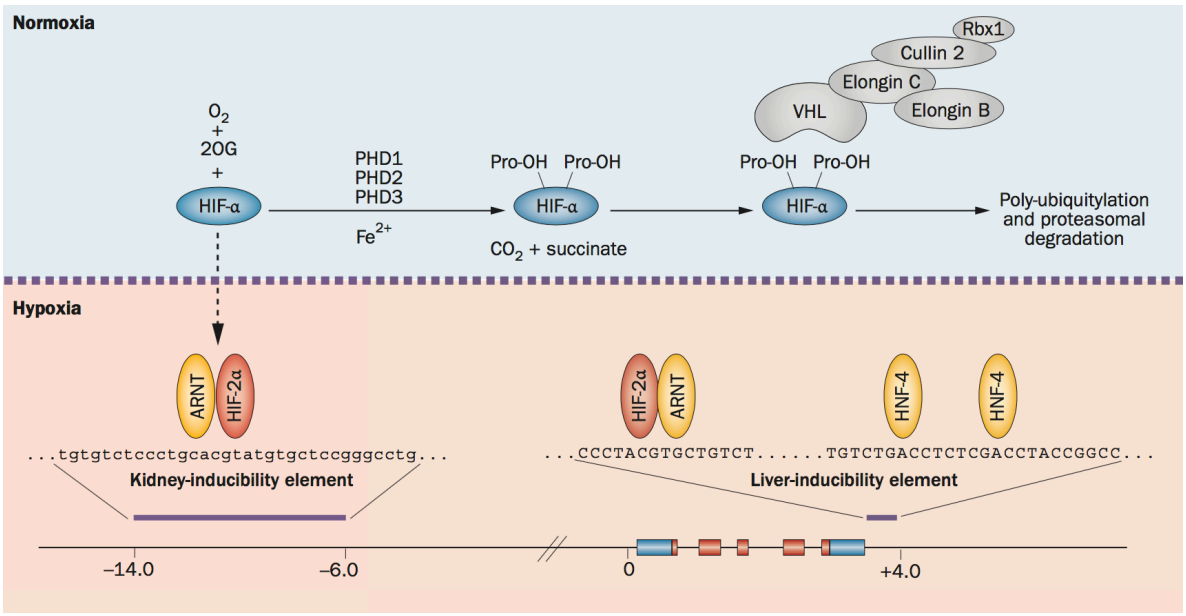


Figure 11. Hypoxic induction of erythropoietin production is regulated by the balance of HIF and PHD proteins. In normoxic conditions, *i.e.* in the presence of oxygen, HIF-α is hydroxylated at specific proline residues by prolyl-4-hydroxylase domain (PHD) oxygenases. Hydroxylation causes the binding of the β-domain of VHL that functions as a substrate recognition component of the VHL-E3-ubiquitin ligase complex. This complex polyubiquitinates HIF-α and targets it for proteasomal degradation (upper panel). Under hypoxic conditions, HIF-α degradation is inhibited, because PHD enzymes need molecular oxygen to hydroxylate HIF-α proteins. HIF-α translocates to the nucleus, where it forms a heterodimer with HIF-ß subunit (ARNT) and binds hypoxia regulatory elements in the proximity of the EPO gene. Liver-inducibility element located downstream of the *EPO* gene mediates hypoxic induction of *EPO* in the liver. Hypoxic induction of renal *EPO* requires the kidney inducibility element located upstream of the *EPO* gene (lower panel.) Boxes depict *EPO* exons. *EPO* coding sequences and non-translated sequences are depicted in red and blue, respectively. The distance from the *EPO* transcription start site is indicated in kilobases. Abbreviations: 2OG, 2-oxoglutarate; ARNT, aryl hydrocarbon receptor nuclear translocator: HNF-4, hepatocyte nuclear factor 4; PHD, prolyl-4-hydroxylase domain. Adapted from (Koury and Haase, 2015).

1.3.5 Dysregulated EPO production and associated clinical syndromes

When renal EPO-producing cells (REPs) lose the ability to produce erythropoietin, it leads to the renal anemia. This is mostly caused by the trans differentiation of REPs into myofibroblasts (Falke et al., 2015; Humphreys et al., 2010; Lin et al., 2008). Patients with renal failure have much lower serum EPO levels when compared to the anemic patients with intact kidney function (Erslev, 1991). Treatment with erythropoiesis-stimulating agents (ESA) such as darbepoetin α or epoetin β leads to the improvements of the renal anemia. Although this therapy is very effective in the majority of patients, some patients do not respond to the treatment. Moreover, varying levels of hemoglobin due to the stimulating agents were associated with increased cardiovascular risk (Besarab et al., 1998; Boudville et al., 2009; Drueke et al., 2006; Pfeffer et al., 2009; Singh et al., 2006). Hence, pharmacological stabilization of HIF proteins may offer an alternative therapeutic

approach with several advantages, such as keeping EPO serum levels within the physiological range, enhanced iron absorption, and the possibility for oral dosing. This stimulated the development of PHD inhibitors (PHIs) (Barrett et al., 2011; Bernhardt et al., 2010; Boettcher et al., 2018; Brigandi et al.; Provenzano et al., 2016). Unlike in the kidney EPCs, where inactivation of PHD2 alone is sufficient to induce erythropoietin expression, all three PHD isoforms have to be inactivated to re-establish *EPO* expression in the liver (Minamishima and Kaelin, 2010).

Also, a genetic mutation in the *EPO* gene resulting in a non-functional protein was reported to cause severe anemia (Kim et al., 2017). Conversely, genetic mutations in *VHL* (Ang et al., 2002), *PHD2* (*EGLN1*) (Ladroue et al., 2008), *HIF2A* (*EPAS2*) (Percy et al., 2008), *HGB* (Charache et al., 1966), *BPGM* (Hoyer et al., 2004), and *EPO* (Zmajkovic et al., 2018) cause excessive production of erythropoietin, leading to the increased red blood cell count and a disease called erythrocytosis, that is discussed in more detail in chapter 1.5.1.

1.4 Thrombopoiesis

Platelets, or thrombocytes, looking like small oval discs with an average size of 2.5x5.0 μm, are essential components of hemostasis and thrombosis. Every day, around 10^{11} platelets are produced by a human body to replace platelets at the end of their 10-day life cycle (Kaushansky, 2005). Platelets are produced from platelet progenitors - megakaryocytes, in a process called platelet biogenesis or megakaryocytopoiesis.

1.4.1 Platelet biogenesis

Burst-colony forming unit megakaryocyte (BFU-MK) cells are the most primitive megakaryocyte-committed progenitors with highest proliferative capacity, that form colonies of more than 50 cells in methylcellulose (Briddell et al., 1989; Long et al., 1985). Next stage is colony-forming unit megakaryocyte progenitor (CFU-MK), with a lower proliferation capacity, that gives rise to colonies of ~ 5 to 50 cells (McLeod et al., 1976) (Vainchenker et al., 1979). CFU-MK progresses through the megakaryoblast stage towards the mature megakaryocyte (Figure 12).

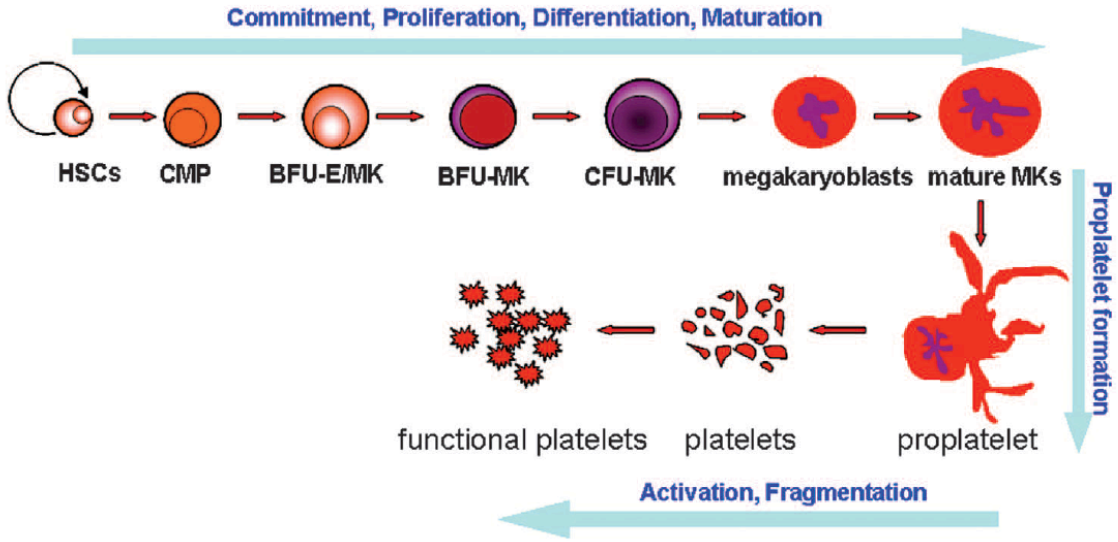


Figure 12. Platelet biogenesis. Different progressive stages of platelet progenitors towards mature thrombocytes are shown. Abbreviations: HSC, hematopoietic stem cell; CMP, common myeloid progenitor; BFU-E/MK, burst-colony forming unit-erythroid/megakaryocyte; BFU-MK, burst-colony forming unit megakaryocyte, CFU-MK, colony-forming unit megakaryocyte; MK, megakaryocyte. Adapted from (Chen et al., 2013).

Megakaryocytes (MK) replicate their DNA without cytokinesis in a process called endomitosis. During endomitosis, multipolar mitotic spindles with the number of poles corresponding to the ploidy level are being assembled, but they fail to separate. This generates a polyploid, multilobed nucleus with up to 128N DNA in man and up to 256N in mouse (Nagata et al., 1997; Vitrat et al., 1998). Besides DNA duplication, also cytoplasmic content undergoes reorganization, and interconnected network of cisternae and tubules called the demarcation membrane system (DMS) is formed (Radley and Haller, 1982). The proplatelet formation is initiated, when megakaryocyte starts to form pseudopodia that continue to elongate until they become proplatelets (Choi et al., 1995; Cramer et al., 1997). Megakaryocytes continue to generate proplatelets until the entire megakaryocyte cytoplasm is transformed into a network of interconnected proplatelets (Italiano et al., 1999). Platelets are formed and released from the swellings at the proplatelet ends (Junt et al., 2007). It was estimated that a single megakaryocyte releases ~ 2000-3000 platelets (Long, 1998).

In addition to thrombopoietin (TPO), a major platelet-stimulating factor, also other growth factors were found to be implicated in platelet biogenesis: IL-3, IL-6 and SCF for early progenitor stage; TGF-β (transforming growth factor β) and IFN-α (interferon α) for megakaryocyte maturation; and BMP4 (bone morphogenic protein 4) and FGF4 (fibroblast growth factor 4) for the platelet release.

1.4.2 Transcriptional regulation of thrombopoiesis

Several transcription factors are involved in the establishment of megakaryocytic lineage. In the early commitment, GATA-1 is essential for lineage commitment towards megakaryocyte-erythrocyte bipotential progenitor, as well as for megakaryocyte development and proliferation (Crispino, 2005). Selective GATA-1 knock-out in megakaryocytic lineage results in thrombocytopenia and higher numbers of immature megakaryocytes with decreased polyploidization (Shivdasani et al., 1997). GATA-1 recruits diverse co-regulators to chromatin, that mediate transcriptional activation or repression (Pope and Bresnick, 2010). One of the examples is Friend of GATA-1 (FOG-1), that binds to GATA-1 and together they synergistically activate transcription from the hematopoietic-specific regulatory region (Tsang et al., 1997). GATA-1 loss-of-function leads to the severe congenital X-linked thrombocytopenia (Nichols et al., 2000) and megakaryoblastic leukemia (Wechsler et al., 2002).

Another transcription factor essential for megakaryopoiesis is Fli-1], regulating late stages of megakaryocytic differentiation (Bastian et al., 1999). Mice null for Fli-1 die during the embryonic development due to the defects in vascular development and megakaryopoiesis Fli-1 co-operates with FOG-1 and GATA-1 to activate the transcription of genes involved in late megakaryopoiesis, such as GPIX or PF4 (Wang et al., 2002).

RUNX-1 (AML-1) binds to the N-terminal part of GATA-1 and contributes to the megakaryocyte lineage commitment (Elagib et al., 2003). Loss-of-function mutations in RUNX1 lead to the familial thrombocytopenia and can progress to AML (Song et al., 1999). The NF-E2 transcription factor, a heterodimeric leucine zipper, was shown to control terminal megakaryocyte maturation, proplatelet formation and platelet release (Deveaux et al., 1997; Lecine et al., 2000). NF-E2 knock-out mice display arrest of megakaryocyte maturation and severe thrombocytopenia, resulting in embryonic lethality and demonstrating the essential role of NF-E2 in platelet biogenesis (Lecine et al., 1998; Shivdasani et al., 1995).

1.4.3 Thrombopoietin (TPO) and thrombopoietin receptor (c-MPL)

Thrombopoietin, also known as a c-Mpl ligand, is the primary cytokine regulating the proliferation of megakaryocytes (Kaushansky et al., 1995). It is produced by hepatocytes in the liver, and to a lesser degree also in the kidney, bone marrow,

and spleen (Sungaran et al., 1997). Besides its role in megakaryopoiesis, it also has a well-documented role in the survival and proliferation of HSCs (Decker et al., 2018; Kaushansky, 2006; Qian et al., 2007; Yoshihara et al., 2007).

The concept of thrombopoietin was first used in 1958 to describe yet unidentified hormone responsible for platelet production (Kelemen et al., 1958). *THPO* gene was not identified until its receptor, *c-MPL* was discovered in 1992 (Skoda et al., 1993; Vigon et al., 1992), by searching for the human and mouse homologs of a viral oncogene *v-MPL* (Souyri et al., 1990). This led to the successful cloning of the *THPO* gene (Bartley et al., 1994; de Sauvage et al., 1994; Kaushansky et al., 1994; Lok et al., 1994). The 6.1 kilobases long human *THPO* is located on chromosome 3q26.3-3q27 and contains 7 exons, of which first 3 are non-coding and have a regulatory function (Foster et al., 1994). Translation of *THPO* mRNA produces a 353-residue polypeptide that upon cleavage of the signal peptide has a length of 332 amino acids and molecular weight of 37.1 kDa. In a long, untranslated 5'-end, 7 upstream open reading frames reside. They are able to initiate the translation and thereby decrease the availability of ribosomes for physiological start codon, AUG8 (Ghilardi et al., 1998). uORF7 extends beyond the AUG8 and is, therefore, the strongest negative regulator of the translation.

Until recently, TPO production was sought to be constitutive (Stoffel et al., 1995), with TPO plasma levels regulated by the uptake of the TPO by megakaryocytes and platelets expressing c-MPL on their surface (Broudy et al., 1997; Fielder et al., 1996; 1997). Hence, TPO concentration in serum would be inversely correlated to the megakaryocyte mass in the bone marrow and platelet counts (Kuter and Rosenberg, 1995; Shinjo et al., 1998). Nevertheless, serum TPO levels are higher than expected in patients with essential thrombocythemia (Griesshammer et al., 1998) and lower than expected in patients with immune thrombocytopenia (Ichikawa et al., 1996; Kosugi et al., 1996). Besides, megakaryocyte-restricted deletion of *c-MPL* and *JAK2* leads to profound megakaryocytosis and thrombocytosis, suggesting that TPO signaling in megakaryocytes is not required for the platelet production (Meyer et al., 2014; Ng et al., 2014). Intriguingly, TPO serum levels in these murine models were in the normal range (Meyer et al., 2014; Ng et al., 2014). Also, IL-6 was shown to increase hepatic *THPO* mRNA expression, as a result of the inflammatory state (Burmester et al., 2005; Kaser et al., 2001; Wolber et al., 2001). In the steady-state conditions, clearance of asialylated senile platelets through the Ashwell-Morell receptor (AMR) was proposed to stimulate *THPO* mRNA expression via

JAK2/STAT3 pathway (Grozovsky et al., 2015). Platelets cleared independently of the AMR do not increase hepatic TPO expression. Furthermore, a recent report attributed a role of glycoprotein receptor GPIbα in hepatic TPO production, indicating that GPIbα glycans on the platelets are recognized by AMR to initiate TPO production (Xu et al., 2018). All this evidence implies that TPO serum levels are regulated in a complex manner, and further investigation is required to fully understand the underlying mechanism.

The gene encoding for the TPO receptor, *c-MPL,* comprises 12 exons located on chromosome 1p34 in man (Mignotte et al., 1994). TPO receptor is a member of the type I cytokine receptor family (Figure 5), and binding of TPO ligand leads to the initiation of JAK2-mediated intracellular signaling events (described in detail in chapter 1.2.1). Knock-out of either *THPO* or *c-MPL* leads to more than 80% reduction in platelet counts and profound thrombocytopenia in mouse models, stressing the critical role of these 2 components in thrombopoiesis (de Sauvage, 1996; Gurney et al., 1994).

Multiple inactivating mutations in *c-MPL* leading to amegakaryocytic thrombocytopenia, a disease that requires hematopoietic stem cell transplantation, were reported (Ballmaier et al., 2001; Ihara et al., 1999; van den Oudenrijn et al., 2000). Recently, also a mutation in the *THPO* gene was reported to cause amegakaryocytic thrombocytopenia (Pecci et al., 2018). Conversely, activating mutations in both thrombopoietin receptor and ligand lead to TPO-independent signaling and cause hereditary thrombocytosis (Cazzola and Skoda, 2000; Ding, 2004) or essential thrombocytosis (Pardanani et al., 2006; Pikman et al., 2006).

For patients suffering from thrombocytopenia, recombinant human TPO was synthesized following the success of rhEPO in treating anemic patients. Despite the initial rise in platelet counts, some of the clinical study subjects developed paradoxical thrombocytopenia, owing to the production of neutralizing autoantibodies (Li et al., 2001). Hence, romiplostim (Bussel et al., 2006; Cwirla et al., 1997) and eltrombopag (Erickson-Miller et al., 2004; Saleh et al., 2013) TPO receptor agonists have been developed as a substitution therapy.

1.5 Hereditary myeloproliferative disorders

Myeloproliferative disorders (MPD) also called myeloproliferative neoplasms (MPN), are a heterogeneous group of clonal blood disorders, characterized by the increased

numbers of mature blood cells. A high level of red blood cells defines polycythemia vera (PV), high level of platelets essential thrombocytosis (ET), and presence of fiber and blasts in bone marrow describes primary myelofibrosis (PMF). These three diseases are distinguished based on blood counts, bone marrow biopsies and the presence of somatic driver mutations in *JAK2* (Baxter et al., 2005; James et al., 2005; Kralovics et al., 2005; Levine et al., 2005), *CALR* (Klampfl et al., 2013; Nangalia et al., 2013), and *c-MPL* (Pikman et al., 2006).

In addition to the spontaneous acquired somatic forms of MPN, there are also familiar forms of MPN existing, usually having clinical symptoms indistinguishable from the somatic forms. Hereditary MPNs are caused by the germline mutations transmitted through Mendelian inheritance within the pedigree. Two classes of hereditary MPNs exist:

1. Inherited disorders with the Mendelian transmission, characterized by high to complete penetrance and polyclonal hematopoiesis.

2. Hereditary predisposition to true MPN, characterized by low penetrance, clonal hematopoiesis, and co-occurrence of somatic mutations.

1.5.1 Hereditary erythrocytosis

Erythrocytosis is diagnosed, when red cell mass is greater than 125% of that predicted for sex and body mass (Keohane et al., 2013). Since determining red cell body mass is a highly specialized test, clinicians use indirect parameters such as hemoglobin and hematocrit values to diagnose erythrocytosis. Based on the most recent WHO guidelines, acquired erythrocytosis (polycythemia vera) is characterized by hemoglobin above 16.5 g/dL in men and 16.0 g/dL in women or hematocrit above 49% in men and 48% in women (Arber et al., 2016).

Primary erythrocytosis is caused by an intrinsic defect in erythroid progenitors or stem cells and is accompanied by low serum EPO levels. Inherited forms of primary erythrocytosis are caused by mutations in the *EPOR* (La Chapelle et al., 1993; Sokol et al., 1995), *SH2B3 (LNK)* (Maslah et al., 2017; Rumi et al., 2016), and *JAK2* (Kapralova et al., 2016).

Secondary erythrocytosis is characterized by elevated EPO serum levels, and is mainly caused by the mutations in genes involved in oxygen sensing pathway, such as *VHL* (Ang et al., 2002; Cario et al., 2005), *PHD2 (EGLN1)* (Ladroue et al., 2008) and *HIF2A (EPAS1)* (Percy et al., 2008). Also, mutations in genes affecting the

hemoglobin *HGB* (Charache et al., 1966), and *BPGM* (Hoyer et al., 2004) were reported to cause erythrocytosis. Most recently we reported a germline mutation in the *EPO* gene to cause secondary erythrocytosis (Zmajkovic et al., 2018). The prevalence of primary erythrocytosis (PV) is estimated to 44-57 cases per 100 000 (Mehta et al., 2014), whereas the prevalence of secondary erythrocytosis is difficult to estimate, but is thought to be < 1: 100 000 (Hussein et al., 2012).

Therapeutic approaches for treating erythrocytosis include phlebotomy (*venesectio)* and low dose aspirin to prevent thrombosis and resulting myocardial infarction and stroke (Landolfi et al., 2004). In addition, several cytoreductive treatments, such as busulfan and hydroxyurea are used, especially for polycythemia vera patients with somatic JAK2 mutations.

1.5.2 Hereditary thrombocytosis

Thrombocytosis, or sometimes called also thrombocythemia is classified by the increase in the number of thrombocytes above $450x10^9$/L for the period over 6 months (Hussein et al., 2014), (Arber et al., 2016). Based on the platelet counts, thrombocytosis may be classified as mild (450-700 $x10^9$/L), moderate (700-900 $x10^9$/L) or severe (>900 $x10^9$/L) (Teofili and Larocca, 2011).

Familial thrombocytosis can be classified as primary when the defect is intrinsic to the megakaryocyte progenitors in the bone marrow. Mutations in the genes involved in the TPO signaling pathway, *c-MPL* (Ding, 2004; El-Harith et al., 2009; Liu et al., 2009; Moliterno et al., 2004; Teofili et al., 2010), and *JAK2* (Bellanne-Chantelot et al., 2006; Mead et al., 2012) lead to the TPO-independent signaling. Primary thrombocytosis is therefore accompanied by the low TPO serum levels. Secondary thrombocytosis is caused by the mutations in the *THPO* gene and is characterized by high TPO serum levels. So far, three distinct mutations in the *THPO* gene were reported as a cause of hereditary thrombocythemia (Ghilardi and Skoda, 1999; Ghilardi et al., 1999; Graziano et al., 2009; Kikuchi et al., 1995; Kondo et al., 1998; Liu et al., 2008; Schlemper et al., 1994; Stockklausner et al., 2012; Wiestner et al., 1998; Zhang et al., 2011). Mutations in other TPO-signaling associated factors, such as *SH2B3*, have not yet been reported. Prevalence of familial thrombocytosis is estimated below 1:100 000 (Teofili and Larocca, 2011). Therapy is based on low aspirin dosing to reduce thromboembolic risk.

2 Results

2.1 A Gain-of-Function Mutation in *EPO* in Familial Erythrocytosis

A Gain-Of-Function Mutation in *EPO* in Familial Erythrocytosis. Zmajkovic J, Lundberg P, Nienhold R, Torgersen ML, Sundan A, Waage A, Skoda RC. N Engl J Med. 378:924-930 (2018).

BRIEF REPORT

A Gain-of-Function Mutation in *EPO* in Familial Erythrocytosis

Jakub Zmajkovic, M.Sc., Pontus Lundberg, Ph.D., Ronny Nienhold, Ph.D., Maria L. Torgersen, Ph.D., Anders Sundan, Ph.D., Anders Waage, M.D., Ph.D., and Radek C. Skoda, M.D.

SUMMARY

Familial erythrocytosis with elevated erythropoietin levels is frequently caused by mutations in genes that regulate oxygen-dependent transcription of the gene encoding erythropoietin (*EPO*). We identified a mutation in *EPO* that cosegregated with disease with a logarithm of the odds (LOD) score of 3.3 in a family with autosomal dominant erythrocytosis. This mutation, a single-nucleotide deletion (c.32delG), introduces a frameshift in exon 2 that interrupts translation of the main *EPO* messenger RNA (mRNA) transcript but initiates excess production of erythropoietin from what is normally a noncoding *EPO* mRNA transcribed from an alternative promoter located in intron 1. (Funded by the Gebert Rüf Foundation and others.)

From the Department of Biomedicine, Experimental Hematology, University Hospital Basel and University of Basel (J.Z., P.L., R.N., R.C.S.), and Diagnostic Hematology, University Hospital Basel (P.L.), Basel, Switzerland; and the Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology (M.L.T., A.S., A.W.), and the Department of Hematology, St. Olavs Hospital (A.W.) — both in Trondheim, Norway. Address reprint requests to Dr. Skoda at radek.skoda@unibas.ch or to Dr. Waage at anders.waage@ntnu.no.

N Engl J Med 2018;378:924-30.
DOI: 10.1056/NEJMoa1709064

Inherited forms of primary erythrocytosis with low serum levels of erythropoietin are caused by mutations in the erythropoietin-receptor gene (*EPOR*),[1,2] *SH2B3* (*LNK*),[3,4] or *JAK2*.[5] Inherited forms of secondary erythrocytosis with elevated erythropoietin levels are mainly caused by mutations in genes involved in oxygen-sensing pathways, such as *VHL*,[6] *PHD2* (*EGLN1*),[7] and *HIF2A* (*EPAS1*),[8] or by mutations in *HGB*[9] or *BPGM*.[10] Variants in candidate genes (e.g., *EGLN2*, *EPO*, *HIF3A*, and *OS9*) have been described,[11] but their functional relevance has not been determined. Here, we describe a gain-of-function variant in *EPO* in an extended kindred with familial erythrocytosis, including 10 affected family members in four generations.

METHODS

STUDY FAMILY

Part of the pedigree — including the index patient, in whom erythrocytosis was diagnosed in 1969 (WG02), and three affected family members (WG01, WG08, and WG11) — has been described previously.[12] The study was approved by the Regional Committee for Medical and Health Research Ethics Central, Norway, and by Ethikkommission Beider Basel, Switzerland. All participating family members provided written informed consent to participate in the study.

GENETIC AND MOLECULAR ANALYSES

Linkage analysis, microsatellite mapping, gene editing with the use of CRISPR (clustered regularly interspaced short palindromic repeats), cell-culture assays, 5′ RACE (rapid amplification of complementary DNA ends), and assessment of erythropoietin levels were performed as described in the Supplementary Appendix, available with the full text of this article at NEJM.org.

RESULTS

CLINICAL FINDINGS AND LABORATORY ANALYSIS

The pedigree of the Norwegian family with autosomal dominant erythrocytosis is shown in Figure 1A. The affected family members typically presented with symptoms of headache and dizziness that rapidly disappeared after initiation of phlebotomy.[12] No incidents of venous or arterial thrombosis had occurred in a family member younger than 70 years of age in this pedigree.

The affected family members had hemoglobin levels above 180 g per liter in men or 160 g per liter in women, with the exception of one male family member (WG05), whose hemoglobin level was in the upper normal range, at 167 g per liter. One child (WG20) had a hemoglobin level of 190 g per liter shortly after birth; we do not have data obtained during infancy for the other members of the family. Additional laboratory data are provided in Table S1 in the Supplementary Appendix. Erythropoietin concentrations in serum were elevated in most of the affected family members. When plotted against the linear regression calculated from hemoglobin and serum erythropoietin levels of healthy persons,[13] the erythropoietin level was found to be elevated in relation to the hemoglobin level in each affected family member (Fig. 1B).

GENOMEWIDE LINKAGE ANALYSIS AND DNA SEQUENCING

Sequence analysis of genes that are known to be involved in erythrocytosis yielded negative results. Therefore, we performed genomewide linkage analysis and found a cosegregating region on chromosome 7q22.1 with a LOD score of 3.3 (Fig. 1C, and Fig. S1 in the Supplementary Appendix). Targeted sequencing of all 215 genes within the cosegregating region (Table S2 in the Supplementary Appendix) revealed a heterozygous single-base deletion in exon 2 of *EPO* (chromosome 7: 100,319,199 GG→G) as the only candidate gene mutation (Fig. 1D). This c.32delG mutation (referred to in this report as ΔG) was present in all affected family members (Fig. 1A) and was absent from 138,632 exome and whole-genome sequences reported in the Genome Aggregation Database (gnomAD).[14] Since erythropoietin is the primary regulator of erythropoiesis, the *EPO* ΔG deletion was an excellent candidate for the disease-causing mutation. The deletion creates a frameshift that truncates the erythropoietin signal peptide and generates a novel peptide, terminating after an additional 51 amino acids (Fig. 1E), which would predict a loss of erythropoietin function and is at odds with the erythrocytosis phenotype.

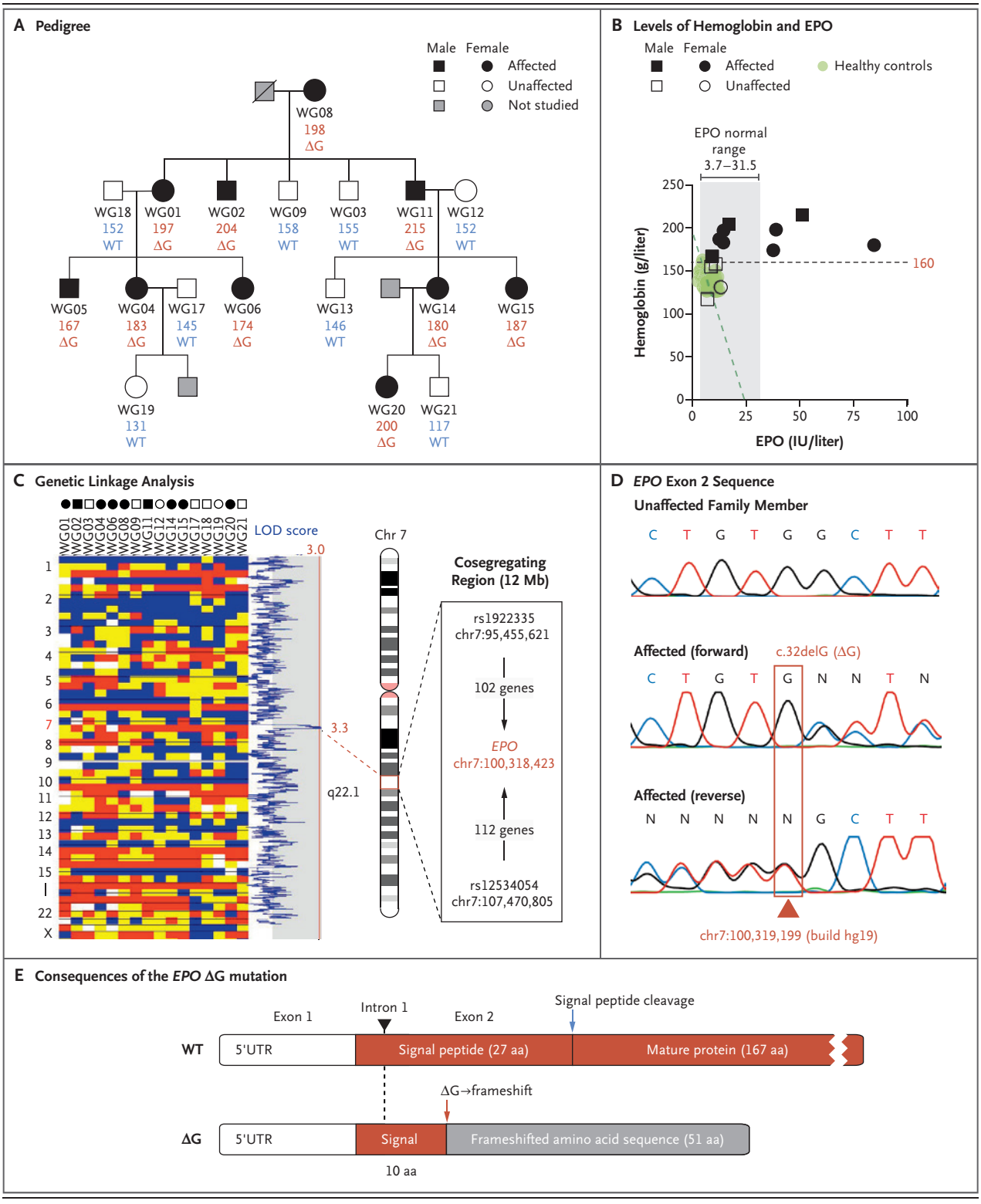
FUNCTIONAL CHARACTERIZATION OF THE MUTATED *EPO*

To determine whether alternative transcripts of *EPO* played a role in the phenotype of these patients, we used CRISPR to introduce the ΔG mutation into *EPO*. Since we had no access to kidney or liver tissues from the patients and were unable to detect *EPO* mRNA in erythroid colonies grown in vitro from a patient's peripheral blood,[15,16] we introduced the mutation into *EPO* in Hep3B cells, a human hepatoma cell line that expresses *EPO*.[17] CRISPR-mediated genome editing (Fig. S2A in the Supplementary Appendix) yielded multiple Hep3B single-cell clones that carried the ΔG mutation in the homozygous state (Fig. 2A, and Fig. S2B in the Supplementary Appendix).

We selected two Hep3B single-cell derived clones (ΔG_1 and ΔG_4) that were homozygous for the *EPO* ΔG mutation and assayed erythropoietin in the supernatants of these clonal cell lines grown in normoxic conditions; the supernatants contained 8 to 10 times as much erythropoietin as the supernatants of parental Hep3B cells or Hep3B cells in which CRISPR failed to modify *EPO* (Fig. 2B). These supernatants were capable of stimulating the growth of an erythropoietin-dependent cell line (BaF3-EpoR), which suggested that the secreted erythropoietin was biologically active (Fig. 2C). The same conclusions were reached with Hep3B cells grown in hypoxic conditions (Fig. S3 in the Supplementary Appendix). We therefore concluded that the ΔG deletion is, in fact, a gain-of-function mutation.

STUDY OF *EPO* MRNA TRANSCRIPTS

To understand how an ostensibly loss-of-function variant of *EPO* produces erythropoietin, we searched for alternative *EPO* mRNAs. In addition to the expected transcripts originating from the physiologic promoter (P1) upstream of exon 1, we discovered two alternative *EPO* mRNA transcripts that originate from an alternative promoter (P2) in intron 1 (Fig. 2D, and Fig. S4A in the Supplementary Appendix). The shorter transcript results from a splicing event that removes 189 nucleotides from the longer intron 1 tran-



script. We detected these two transcripts in parental Hep3B cells as well as in the ΔG clones; the longer P2 transcript was more abundant in the ΔG_1 and ΔG_4 clones than in parental Hep3B cells when grown in either normoxic (Fig. 2E) or hypoxic conditions. P1 and P2 transcripts were also detected in mRNA from normal kidney and liver (Fig. 2F). In kidney, the P1 transcript was

Figure 1 (facing page). Pedigree of the Family with Hereditary Erythrocytosis and Identification of the Mutation in *EPO*, the Gene Encoding Erythropoietin.

Panel A shows a pedigree of the family in this study. Unique patient numbers are placed under the symbols, with numbers below representing hemoglobin levels in grams per liter and the *EPO* sequence status indicated as wild-type (WT) or as having the c.32delG mutation (ΔG). Panel B shows hemoglobin levels plotted against erythropoietin (EPO) levels in serum. The green circles represent hemoglobin and EPO levels of 35 healthy persons,[13] and the green dashed line is the linear regression calculated from these values and represents the expected normal EPO level for a given hemoglobin level. The black dashed line indicates the hemoglobin level of 160 g per liter, which is the upper limit of the normal range among females. The gray shaded area indicates the normal range of EPO levels in serum. Panel C shows mapping of the disease-causing gene locus by the Affymetrix 250K single-nucleotide polymorphism (SNP) array. Red, yellow, and blue boxes indicate genotype calls for each SNP in the array (with alleles at each SNP arbitrarily designated as A and B, red denotes AA, yellow AB, and blue BB). The vertical red line indicates a logarithm of the odds (LOD) score of 3.0. An enlarged view of the cosegregating region on chromosome 7 is shown on the right. The nucleotide positions of the two SNPs that mark the borders of the cosegregating region are also shown. Panel D shows chromatograms of the DNA sequences of the mutated locus. The heterozygous loss of one G nucleotide results in ambiguous sequence reads downstream or upstream of the deleted G in the forward and reverse chromatograms, respectively. The red box indicates the location of the single-base deletion (ΔG). Panel E shows the location of the ΔG mutation and the resulting frameshift in the *EPO* mRNA. The wild-type EPO coding sequence is shown in red, and the sequences encoding the signal peptide and the mature protein are marked. The ΔG mutation induces a frameshift in the *EPO* mRNA sequence that truncates the signal peptide and generates a novel peptide, terminating after additional 51 amino acids (gray box). UTR denotes untranslated region.

predominant, whereas in the liver, the long P2 mRNA was predominant. In Hep3B cells, the presence of the ΔG mutation did not alter the rate of transcription from the P1 and P2 promoters. However, the ΔG P2 mRNAs had a longer half-life than did the wild-type P2 mRNAs, which suggested that increased mRNA stability is responsible for the elevated levels of expression of P2 transcripts observed in Hep3B cells with the ΔG mutation (Fig. 2E). Additional information on the characterization of the P1 and P2 transcripts is provided in Figure S4 in the Supplementary Appendix.

To determine which transcripts are capable of producing erythropoietin, we transfected HEK293 cells with complementary DNAs (cDNAs) representing P1 or P2 transcripts with or without the ΔG mutation and measured erythropoietin in culture supernatants. Cells transfected with ΔG P2 transcripts produced more erythropoietin than cells transfected with wild-type P1 cDNA (Fig. S5A in the Supplementary Appendix). These supernatants also stimulated the growth of erythropoietin-dependent BaF3-EpoR cells (Fig. S5B in the Supplementary Appendix) and supported the growth of erythroid colonies from human peripheral blood (not shown), which showed that ΔG P2 transcripts produce biologically active erythropoietin.

TRANSLATION OF *EPO* MRNA VARIANTS

We performed in vitro transcription–translation experiments and examined the composition of the open reading frames (ORFs) in *EPO* mRNAs to determine why more erythropoietin was produced by the short ΔG P2 cDNA construct than by the long ΔG P2 construct in transfected HEK293 cells (Fig. S5 in the Supplementary Appendix). In vitro translation of the short ΔG P2 transcript produced more erythropoietin than did translation of the long ΔG P2 transcript or the wild-type P1 mRNA because the splicing event eliminates most of the upstream ORFs, which inhibit translation. Since AUG1⋆ (the first start codon in the P2 transcripts; asterisks are used here to highlight start codons located in intron 1) is located within a sequence context that is predicted to be weak at initiating translation (Table S3 in the Supplementary Appendix),[18] ribosomes are likely to skip AUG1⋆ and initiate translation from the next available start codon, which, in the short ΔG P2 transcript, is AUG2, located in exon 2. In vitro translation of the long ΔG P2 mRNA, most likely initiating from upstream AUG3⋆ located in intron 1, produced a larger erythropoietin molecule that was approximately 25 kDa in size. The smaller observed protein product, which is approximately 21.6 kDa in size, is consistent with initiation of translation from AUG2. Removing upstream ORF1 by deleting the 5′ untranslated region (ΔUTR) increased the translational efficiency of the wild-type P1 mRNA, as reported previously.[19] HEK293 cells expressing the ΔG P1 cDNA produced low amounts of erythropoietin. A likely explanation is that ribosomes can skip the physiologic start codon (AUG1) in the ΔG P1 mRNA and initiate translation from AUG2, which connects with the erythropoietin coding sequence through the ΔG

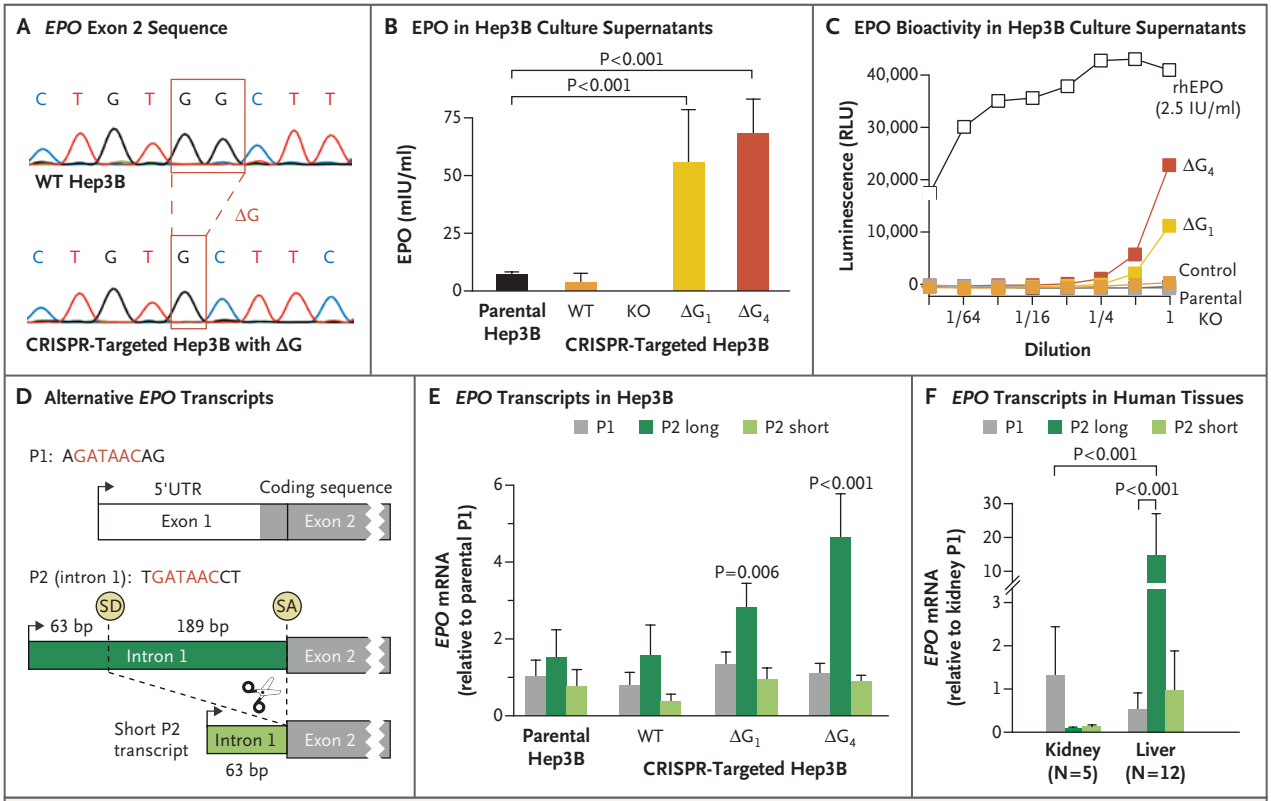


Figure 2. Functional Characterization of the *EPO* ΔG Mutation.

Panel A shows chromatograms of *EPO* exon 2 DNA sequences from WT Hep3B and a CRISPR-modified single-cell clone. The region containing the ΔG deletion is indicated by the red box. Panel B shows concentrations of EPO in culture supernatants from Hep3B cells carrying the ΔG mutation. Parental Hep3B cells, Hep3B cells in which CRISPR failed to modify the *EPO* gene (WT), and Hep3B cells with *EPO* inactivated by a deletion and insertion (KO) are shown for comparison. EPO was measured by means of enzyme-linked immunosorbent assay (ELISA), and EPO concentrations are given as means and standard deviations (T bars) of three biologic replicates. P values were determined by one-way analysis of variance. Panel C shows the biologic activity of EPO produced by parental Hep3B or CRISPR-modified Hep3B cells, measured in culture supernatants by a proliferation assay with the BaF3-EpoR cell line. Serial dilutions of 2.5 IU per milliliter of recombinant human EPO (rhEPO) were used for reference (based on three biologic replicates). Panel D shows the *EPO* transcripts identified by 5′-RACE in Hep3B cells. The P1 transcript starts at exon 1, and its transcription is initiated from *EPO* promoter 1 (P1). A putative promoter 2 (P2) is located in *EPO* intron 1. This promoter generates an alternative *EPO* transcript P2 (dark green) that can be further spliced to produce a short version of the P2 transcript (light green). Splice donor (SD) and splice acceptor (SA) sites are indicated with dashed lines. The GATA motif in the putative promoter is shown in red. The 5′ untranslated region (UTR) and the protein-coding region (gray shading) are shown for the P1 transcript. Panel E shows quantification of *EPO* transcripts in CRISPR-modified and parental Hep3B cell lines by means of quantitative reverse-transcriptase polymerase chain reaction (RT-qPCR). Values are the means and standard deviations (T bars) of four biologic replicates. of mRNA expression relative to expression of the parental P1 transcript, which was defined as 1. (For example, a value of 3 indicates expression 3 times as high as that of the parental P1 transcript, and a value of 0.5 indicates expression half as high.) P values were determined by two-way analysis of variance. Panel F shows quantification by RT-qPCR of *EPO* transcripts in human kidney and liver RNA. Values represent means and standard deviations (T bars) of mRNA expression relative to expression of the P1 transcript in kidney, which was defined as 1. P values were determined by two-way analysis of variance.

frameshift. Thus, erythropoietin protein from all ΔG transcripts appears to be produced by initiating translation from AUG2. Additional information on erythropoietin production by the P1 and P2 transcripts is provided in Figure S5 in the Supplementary Appendix.

Initiation of translation from AUG2 produces an erythropoietin protein with a shortened signal peptide (22 instead of 27 amino acids) and a novel N-terminal, which nevertheless was predicted by computer-based algorithm (SignalP 4.1[20]) to be functional (Fig. S6 in the Supplementary Appendix). To test this prediction, we generated a cDNA construct representing *EPO* ΔG that starts directly with AUG2 and lacks any upstream AUGs. This construct (P2 ΔG ΔUTR) was most active in producing erythropoietin in vitro and in transfected HEK293 cells; it also showed bio-

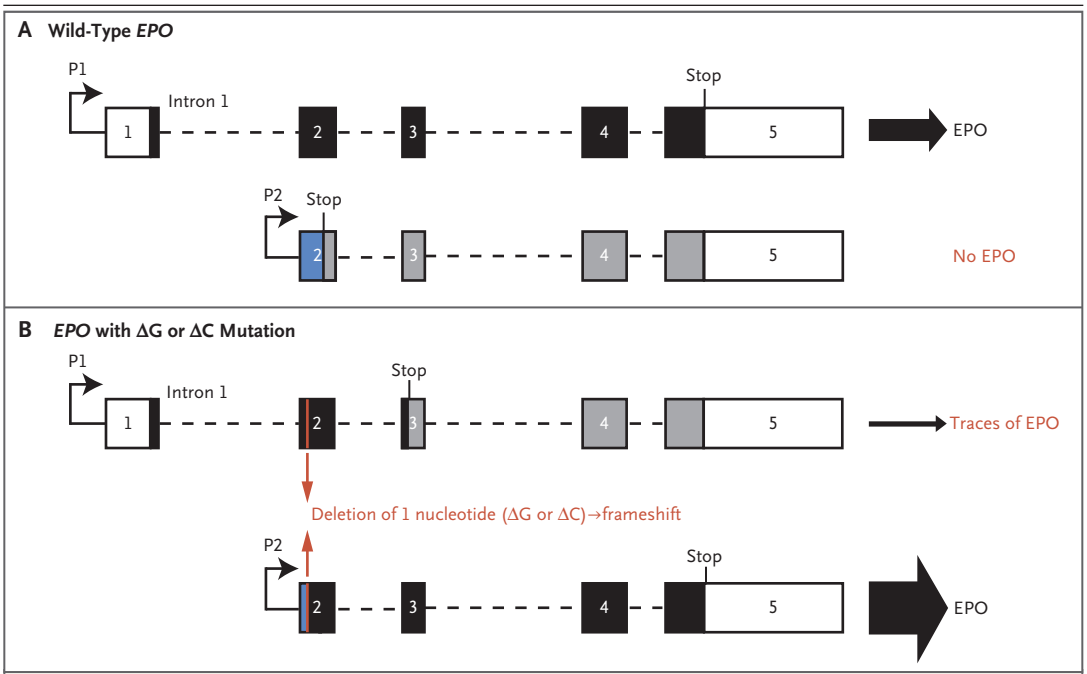


Figure 3. Model and Proposed Mechanism for EPO Overproduction by a Single-Nucleotide Deletion in Exon 2 of *EPO*.

The exons of *EPO* are numbered and drawn as boxes. Open boxes represent 5′ and 3′ UTRs. Panel A shows the wild-type gene producing EPO from P1 transcripts (black boxes), whereas P2 promoter transcripts generate only a short, nonfunctional peptide (blue box). Panel B shows the effects of the frameshift caused by a single-nucleotide deletion in exon 2 of *EPO* (ΔG or ΔC). The mutated P1 transcripts produce only a short nonfunctional peptide terminating in exon 3 (black boxes), and minimal amounts of EPO can be produced if translation is initiated from an alternative AUG2 start site (see Fig. S5A in the Supplementary Appendix). P2 transcripts (blue box) connect through the frameshift to the EPO coding sequence (black boxes) and are responsible for the overproduction of EPO. Gray boxes represent EPO coding sequences that are in a different reading frame.

logic activity in the BaF3-EpoR proliferation assay (Fig. S5A and S5B in the Supplementary Appendix), which indicated that the shortened signal peptide is indeed functional.

SECOND *EPO* ALLELE ASSOCIATED WITH ERYTHROCYTOSIS

A father and daughter with erythrocytosis and a single C nucleotide deletion (ΔC) of unknown significance, which was located in *EPO* exon 2 just 13 nucleotides upstream of ΔG (c.19delC; chromosome 7: 100,319,185 TC→T), were recently described.[11] To test whether *EPO* ΔC has the same effect as *EPO* ΔG, we generated ΔC P2 cDNA constructs, transfected HEK293 cells, and measured erythropoietin protein and erythropoietin bioactivity in cell-culture supernatants. The long P2 and short P2 transcripts with the ΔC mutation also produced high amounts of erythropoietin that was biologically active (Fig. S7 in the Supplementary Appendix), which showed that the *EPO* ΔC mutation causes erythrocytosis through the same mechanism (Fig. 3) as the *EPO* ΔG mutation.

DISCUSSION

We propose that the frameshift mutation ΔG in *EPO* causes erythrocytosis on the basis of genetic linkage with a locus on chromosome 7q21 (LOD score, 3.3), functional data, and our finding of a second, independent allele (ΔC) in a father and daughter with erythrocytosis. Both ΔG and ΔC cause a frameshift and convert a noncoding mRNA transcribed from an alternative (P2) promoter inside *EPO* intron 1 into an mRNA that produces functional erythropoietin protein. This frameshift occurs in the sequence encoding the signal peptide of erythropoietin; the mutant peptide sequence is fully functional. We dissected the relative contribution of individual upstream AUG start codons located in the 5′ untranslated regions of P1 and P2 transcripts and established that erythropoietin protein from both P2 ΔG transcripts appears to be produced by initiating translation from AUG2, located in exon 2.

P2 transcripts were more abundant in normal liver than in kidney, which suggested that the

main source of erythropoietin production in affected family members was the liver. Indeed, P2 transcripts are present in human liver RNAseq data sets from ProteinAtlas[21] and the Illumina Human Body Map (Fig. S8 in the Supplementary Appendix). In a database of capped 5′ ends of mRNAs,[22,23] P2 transcripts were the predominant mRNA species in the liver, whereas P1 transcripts were mainly detected in RNA from kidney (Fig. S9 in the Supplementary Appendix). The physiological function of wild-type P2 transcripts is currently unknown, but it seems likely that they do have a function, because *EPO* intron 1 has a high degree of sequence homology between human and mouse and the region of homologous sequence extends upstream of the P2 promotor (Fig. S10 in the Supplementary Appendix).[24]

Our data indicate that the *EPO* ΔG mutation prematurely terminates translation of *EPO* P1 transcripts, whereas it alters the normally noncoding P2 transcripts such that AUG2 initiates translation of the erythropoietin coding sequence and produces an excess of biologically active erythropoietin. Mutations in *EPO* should be considered in the search for causes of inherited erythrocytosis. More generally, the effect of mutations on multiple mRNA transcripts may explain nonintuitive relationships between phenotype and the ostensible mechanism predicted by the mutation.

Supported by grants from the Gebert Rüf Foundation (GRS-045/09), the Swiss Cancer League (KLS-2950-02-2012 and KFS-3655-02-2015), and the Swiss National Science Foundation (31003A-147016/1 and 31003A_166613) (all to Dr. Skoda); a University of Basel "Get on Track" scholarship (to Mr. Zmajkovic); and a grant from the Central Norway Regional Health Authority (Biobank Research 2011) (to Dr. Waage).

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Mihaela Zavolan and members of the laboratory for helpful discussions and critical reading of an earlier version of the manuscript; Oliver Mühlemann for advice on RNA stability and nonsense mediated decay; Helen A. Papadaki for providing hemoglobin and erythropoietin levels of healthy persons; Hélène Méreau for help with the microscopy; Nitish Mittal and Florian Groelly for help with the in vitro translation; Florian Geier and Foivos Gypas for help with the bioinformatics analysis; Philippe Demougin, Christian Beisel, and the staff of Genomics Facility Basel for conducting next-generation sequencing; Lill Anny Grøseth and Solveig Kvam for work with the samples collected in the Biobank for hematologic diseases (Biobank1) in St. Olavs Hospital, Trondheim, Norway; and the family members for their cooperation.

REFERENCES

1. de la Chapelle A, Träskelin AL, Juvonen E. Truncated erythropoietin receptor causes dominantly inherited benign human erythrocytosis. Proc Natl Acad Sci U S A 1993;90:4495-9.
2. Sokol L, Luhovy M, Guan Y, Prchal JF, Semenza GL, Prchal JT. Primary familial polycythemia: a frameshift mutation in the erythropoietin receptor gene and increased sensitivity of erythroid progenitors to erythropoietin. Blood 1995;86:15-22.
3. Rumi E, Harutyunyan AS, Pietra D, et al. LNK mutations in familial myeloproliferative neoplasms. Blood 2016;128:144-5.
4. Maslah N, Cassinat B, Verger E, Kiladjian JJ, Velazquez L. The role of LNK/SH2B3 genetic alterations in myeloproliferative neoplasms and other hematological disorders. Leukemia 2017;31:1661-70.
5. Kapralova K, Horvathova M, Pecquet C, et al. Cooperation of germ line JAK2 mutations E846D and R1063H in hereditary erythrocytosis with megakaryocytic atypia. Blood 2016;128:1418-23.
6. Ang SO, Chen H, Hirota K, et al. Disruption of oxygen homeostasis underlies congenital Chuvash polycythemia. Nat Genet 2002;32:614-21.
7. Ladroue C, Carcenac R, Leporrier M, et al. PHD2 mutation and congenital erythrocytosis with paraganglioma. N Engl J Med 2008;359:2685-92.
8. Percy MJ, Furlow PW, Lucas GS, et al. A gain-of-function mutation in the HIF2A gene in familial erythrocytosis. N Engl J Med 2008;358:162-8.
9. Charache S, Weatherall DJ, Clegg JB. Polycythemia associated with a hemoglobinopathy. J Clin Invest 1966;45:813-22.
10. Hoyer JD, Allen SL, Beutler E, Kubik K, West C, Fairbanks VF. Erythrocytosis due to bisphosphoglycerate mutase deficiency with concurrent glucose-6-phosphate dehydrogenase (G-6-PD) deficiency. Am J Hematol 2004;75:205-8.
11. Camps C, Petousi N, Bento C, et al. Gene panel sequencing improves the diagnostic work-up of patients with idiopathic erythrocytosis and identifies new mutations. Haematologica 2016;101:1306-18.
12. Ly B, Meberg A, Kannelønning K, Refsum HE, Berg K. Dominant familial erythrocytosis with low plasma erythropoitin activity: studies on four cases. Scand J Haematol 1983;31:11-7.
13. Papadaki HA, Kritikos HD, Valatas V, Boumpas DT, Eliopoulos GD. Anemia of chronic disease in rheumatoid arthritis is associated with increased apoptosis of bone marrow erythroid cells: improvement following anti-tumor necrosis factor-alpha antibody therapy. Blood 2002;100:474-82.
14. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285-91.
15. Stopka T, Zivny JH, Stopkova P, Prchal JF, Prchal JT. Human hematopoietic progenitors express erythropoietin. Blood 1998;91:3766-72.
16. Sato T, Maekawa T, Watanabe S, Tsuji K, Nakahata T. Erythroid progenitors differentiate and mature in response to endogenous erythropoietin. J Clin Invest 2000;106:263-70.
17. Goldberg MA, Glass GA, Cunningham JM, Bunn HF. The regulated expression of erythropoietin by two human hepatoma cell lines. Proc Natl Acad Sci U S A 1987;84:7972-6.
18. Kozak M. Interpreting cDNA sequences: some insights from studies on translation. Mamm Genome 1996;7:563-74.
19. Barbosa C, Romão L. Translation of the human erythropoietin transcript is regulated by an upstream open reading frame in response to hypoxia. RNA 2014;20:594-608.
20. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 2011;8:785-6.
21. Uhlén M, Fagerberg L, Hallström BM, et al. Proteomics: tissue-based map of the human proteome. Science 2015;347:1260419.
22. Forrest AR, Kawaji H, Rehli M, et al. A promoter-level mammalian expression atlas. Nature 2014;507:462-70.
23. Lizio M, Harshbarger J, Shimoji H, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol 2015;16:22.
24. Shoemaker CB, Mitsock LD. Murine erythropoietin gene: cloning, expression, and human gene homology. Mol Cell Biol 1986;6:849-58.

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Zmajkovic J, Lundberg P, Nienhold R, et al. A gain-of-function mutation in *EPO* in familial erythrocytosis. N Engl J Med 2018;378:924-30. DOI: 10.1056/NEJMoa1709064

SUPPLEMENTARY APPENDIX

Supplement to: Zmajkovic J, Lundberg P, Nienhold R, Torgersen Lyngaas M, Sundan A, Waage A and Skoda RC: A-gain-of-function mutation in *EPO* gene causes familial erythrocytosis.

This appendix has been provided by the authors to give readers additional information about their work.

Contents **Page**

Supplementary methods

Patient samples and measurements of blood parameters

Measurements of Hb, leukocytes, platelets and Epo were carried out at the Department of Medical Biochemistry, St. Olavs Hospital, Trondheim, Norway during the time period of 1966-2017. 80 % of the measurements have been carried out during a time when the normal range of Hb has been 117 - 153 g/L in women and 134 - 165 or 170 g/L in Men. The normal range for Epo in serum was 3.7 - 31.5 IU/L.

Linkage analysis and next generation sequencing

We performed parametric linkage analysis using Affymetrix SNP 250K arrays, and dChip software[1] for linkage analysis. Next generation sequencing was performed using Agilent SureSelect Target Enrichment designed to capture exons and intron/exon boundaries of co-segregating regions. CLC Genomic Workbench was used for alignment and variant detection comparing against the human reference genome build hg19.

Microsatellite mapping

10 FAM-labelled microsatellite markers (sequences in Supplementary Table 4) spanning from q21.11 (D7S669) to q32.2 (D7S530) on chromosome 7 were used for genotyping of 18 members of the family. PCR was performed with Q5 High-Fidelity DNA Polymerase (NEB) and subsequently analysed on an ABI 3130xl Genetic Analyzer (Applied Biosystems). Data collection and allele identification were performed using GeneMapper v5.0 software (Applied Biosystems).

EPO gene sequence analysis

Primers flanking exon 2 of human *EPO* (fwd: 5'-GCCACCACTTATCTGCCA-3', rev: 5'-AACTCCACCCCAAACCAA-3') were used for PCR amplification with Q5 High-Fidelity DNA Polymerase (NEB). PCR products were purified with Exonuclease I and Thermosensitive Alkaline Phosphatase mix (Thermo Fisher) by incubating 15min. at 37°C followed by 85°C for 15min. Reactions were submitted for Sanger sequencing to Microsynth. Sequences were analysed with CLC Main Workbench 7.7.1 (Qiagen).

Cell lines and transfections

Human hepatocellular carcinoma cell line Hep3B (kindly provided by Drs. Luca Quagliata and Cristina Quintavalle from Institute of Pathology, University Hospital Basel) was maintained in RPMI media with L-glutamine and HEPES (GIBCO Invitrogen) supplemented with 10% fetal bovine serum (Sigma) and Penicillin-Streptomycin mix (GIBCO Invitrogen). Genomic DNA was isolated with QIAamp DNA mini kit (Qiagen) and sent for STR genotyping to Microsynth that

confirmed Hep3B identity and excluded contamination with other human cell lines. HEK293LTV cell line (Cell Biolabs) was maintained in high glucose DMEM with L-glutamine supplemented with 10% fetal bovine serum (Sigma) and Penicillin/Streptomycin mix (GIBCO Invitrogen). BaF3-EpoR cells were cultured in RPMI media same as Hep3B with 10% conditioned medium of WEHI-3 cell line. Cells were grown at 37 °C in 5 % CO_2 humidified atmosphere with 21% O_2. For hypoxia experiments, Hep3B cells were shifted to 1% O_2, 5% CO_2 for 72 hours, if not stated otherwise. For overexpression experiments, $2{,}5x10^5$ cells per well (6-well plates) were transfected with 0.5 µg of the plasmid using jetPRIME reagent (Polyplus Transfection) and supernatants were harvested 72 hours post transfection.

Single-guide RNA (sgRNA) cloning

Single-guide RNA targeting exon 2 of *EPO* was designed with Zhang algorithm (crispr.mit.edu) and cloned into pSpCas9(BB)-2A-GFP (pX458) vector (obtained from Addgene) using primers in Supplementary Table 4 and following previously established protocol.[2]

CRISPR-mediated *EPO* locus targeting

5 µg of targeting vector was premixed with 100 pmoles of 150 bp single-stranded oligo donor template (Integrated DNA Technologies) and electroporated into $1x10^6$ Hep3B cells per condition using Cell Line Nucleofactor Kit V (Lonza). GFP-positive cells were sorted 48 hours post nucleofection as a bulk and expanded for 1 week in culture. After expansion, cells were divided into pools and genomic DNA was isolated using QIAamp DNA mini kit (Qiagen). To assess, which pools contain the highest frequency of mutated cells, SNP genotyping assay (Applied Biosystems) was performed using custom primers and TaqMan probes (sequences in Supplementary Table 4). SNP genotyping PCR was performed on ABI 7500 Fast Real-Time PCR System (Applied Biosystems) with 15ng of genomic DNA using 2x TaqMan Universal PCR Master Mix (Applied Biosystems).

Cell pools with highest frequency of mutated cells were single-cell sorted into 96-well plates coated with Collagen Solution (StemCell Technologies) and cultured in 1:1 fresh RPMI media and Hep3B-conditioned RPMI media. Genomic DNA was isolated by incubating cells in DNA lysis buffer (50mM KCl, 10mM Tris pH8, 0.45% Tween-20, 0.45% NP-40 and 20mg/ml proteinase K) at 65°C for 60min. followed by 10min. at 95°C. 5µl of mixture were used for *EPO* exon 2 PCR followed by Sanger sequencing as described above.

To confirm both alleles of *EPO* were successfully targeted, PCR products were subcloned with Zero Blunt PCR Cloning Kit (Invitrogen), transformed into the NEB 5-alpha Competent E.coli (NEB), plasmid DNA was isolated using NucleoSpin Plasmid kit (Macherey-Nagel) and send out for sequencing to Microsynth. Sequences were analysed with CLC Main Workbench 7.7.1 (Qiagen).

CRISPR off-target analysis

CRISPR-edited Hep3B clones were tested for off-target editing events predicted for sgRNA by the Zhang laboratory algorithm (crispr.mit.edu) and the COSMID tool [3] (crispr.bme.gatech.edu). The top-five non-overlapping predicted off-target sites were used from each tool (Supplementary Table 5). Regions surrounding each off-target site were PCR-amplified using primers in Supplementary Table 4, Sanger sequenced as described above and compared to the unedited parental cell line.

Epo levels measurement

Cell supernatants from Hep3B cell lines cultured under normoxia and hypoxia were sterile-filtered (0.22μm) and Epo levels were measured with human erythropoietin Legend Max ELISA Kit (Biolegend). Epo levels were normalized to the protein concentration in the cell lysate. Cell supernatants from transfected HEK293LTV cells were sterile-filtered (0.22μm) and Epo levels were measured with Human EPO Tissue Culture Kit (Meso Scale Discovery) on MESO Sector S 600 device equipped with Discovery Workbench 4.0 software.

Proliferation assay

Cell proliferation rates were measured by using CellTiter-Glo Luminiscent Cell Viability Assay (Promega). Supernatants from Hep3B cell lines and transfected HEK293LTV cells were used to stimulate the growth of BaF3-EpoR cell line. Recombinant human Epo (Peprotech) was used as a positive control. BaF3-EpoR cells were washed 3 times with PBS and $5x10^3$ cells per well were plated into 96-well plates and incubated in RPMI + 10% FBS with either 2- or 10-fold dilution of cell supernatants or recombinant human Epo for 72 hours. Then 30μl of cells were transferred into black opaque 96-well plates (Eppendorf) and incubated with equal volume of Cell-Titer-Glo Reagent according to the manufacturer's protocol. Luminiscence was measured on Synergy H1 microplate reader (Biotek) equipped with Gen5 v2.05 software.

5' rapid amplification of cDNA ends (5'RACE)

Total RNA from parental and CRISPR-edited Hep3B cell lines isolated with TriFast (VWR International) together with total RNA from human kidney and liver (Takara Clontech) were subjected to 5'RACE with SMARTer RACE 5'/3' Kit (Takara Clontech) according to the manufacturer's protocol with the following gene specific primers for human *EPO*: GSP1: 5'-GATTACGCCAAGCTT**AGCAGTGATTGTTCGGAGTGGAGCAG** and GSP 2-GATTACGCCAAGCTT**AGAGCCCGAAGCAGAGTGGTGAGGC**, gene-specific part in bold. RACE products were further subcloned and transformed, plasmid DNA from bacterial colonies was isolated using Zyppy-96 Plasmid Miniprep (Zymo Research) and Sanger sequenced. The sequencing reads were mapped to reference gene variant with CLC Main Workbench 7.7.1 (Qiagen).

Cloning of *EPO* constructs and site-directed mutagenesis

EPO transcripts were amplified from Hep3B cDNA with primers listed in Supplementary Table 4 with Q5 High-Fidelity DNA Polymerase mix (NEB). PCR products were purified from agarose gel with NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) and subsequently cloned into pcDNA3 vector (Invitrogen). For introducing ΔG and ΔC mutations into the plasmids, QuickChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies) was used with primers listed in Supplementary Table 4. Plasmid sequences were verified with Sanger sequencing and midiprep DNA was produced with NucleoBond Xtra Midi kit (Macherey-Nagel).

***EPO* transcripts quantitation**

Total RNA from parental and CRISPR-edited Hep3B cell lines isolated with TriFast (VWR International) and together with total RNA from human kidney and liver (Takara Clontech and AMS Biotechnology) was reverse-transcribed with High Capacity cDNA Reverse Transcription Kits (Applied Biosystems). cDNA was amplified with Power SYBR Green PCR Master Mix (Applied Biosystems) and primers listed in Supplementary Table 4 on ViiA 7 Real-Time PCR System (Applied Biosystems). Relative gene expression change was calculated according to ΔΔCt method using *GUSB* gene for normalization.

RNA transcription rate

To investigate RNA transcription rates, Hep3B cells were plated at density $1{,}5x10^5$ cells per well (12-well plates) and 24 hours later, 0.5mM of 5-ethynyl uridine (5-EU) was added to the cells with fresh media for 30min. Nascent RNAs were captured by using the Click-iT Nascent RNA Capture Kit (Invitrogen) in accordance with the manufacturer's instructions.

RNA stability

Stability of *EPO* mRNAs was examined as described previously.[4] Briefly, Hep3B cells were plated at density $2{,}5x10^5$ cells per well (6-well plates) and first incubated in hypoxic chamber ($1\%O_2$, 5% CO_2) for 48 hours and then switched to the normoxia ($21\%O_2$, 5% CO_2). Cells were collected at 0, 1, 2, 3, 4, 6 and 8 hrs after the switch to normoxia. RNA isolation, cDNA synthesis and qRT-PCR was performed as described above.

***In vitro* transcription and translation**

EPO constructs in pcDNA3 vector were linearized with XhoI, extracted with phenol-chloroform and ethanol-precipitated. 1µg of linearized DNA was used as a template for *in vitro* transcription using T7 MEGAScript kit (Thermo Fisher Scientific). Reactions were incubated for 4 hrs at 37°C, followed by DNase treatment and phenol-chloroform extraction. Identity of the different *EPO* transcripts was verified by Agilent 2100 Bioanalyzer. 2.5ug of each *EPO* mRNA was translated for 2hrs at 25°C using wheat germ extract lysate

(Promega) in the presence of FluoroTect Green *in vitro* Translation Labeling System (Promega). Labelled proteins were separated on NuPAGE 12% Bis-Tris Protein gels (Invitrogen) and visualized on Typhoon 9400 Imaging System (GE Healthcare Life Sciences) using 488nm laser and 526SP emission filter.

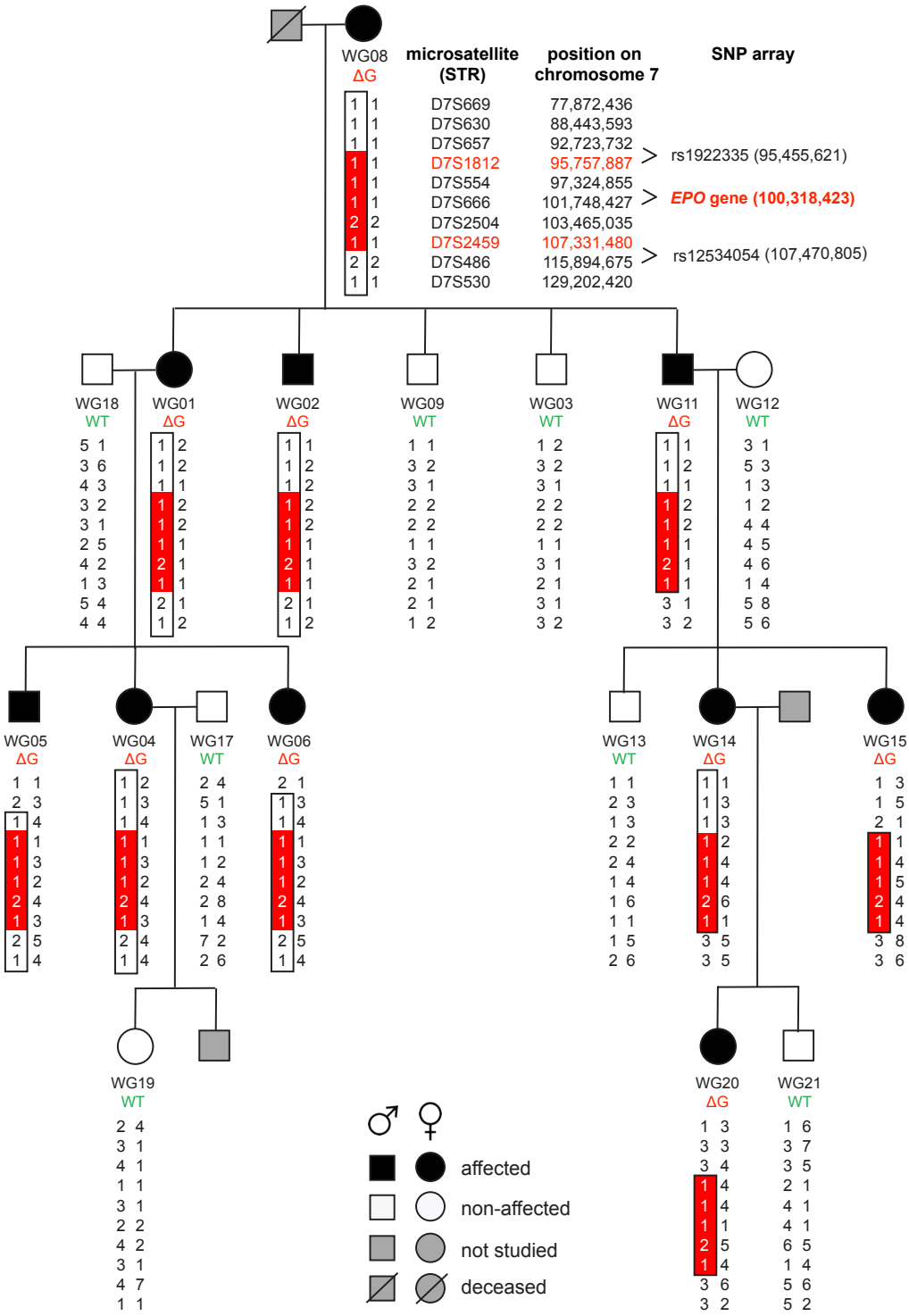
Bioinformatics analysis

RNAseq reads were downloaded from the Human Protein Atlas [5] or Illumina Body Map, 3' adapters were trimmed using Cutadapt. [6] The trimmed reads were mapped (only mate 1 was used) to the human genome using the STAR aligner (version 2.5.3a) with default parameters. [7] The ENSEMBL (version 87) annotation file was used. [8] The mapped reads were then sorted and indexed using SAMtools (version 1.3). [9] Processed and aligned CAGE-sequencing data were downloaded from FANTOM5 project. [10,11] Read coverage was generated with IGV v2.3. [12]

Statistical analysis

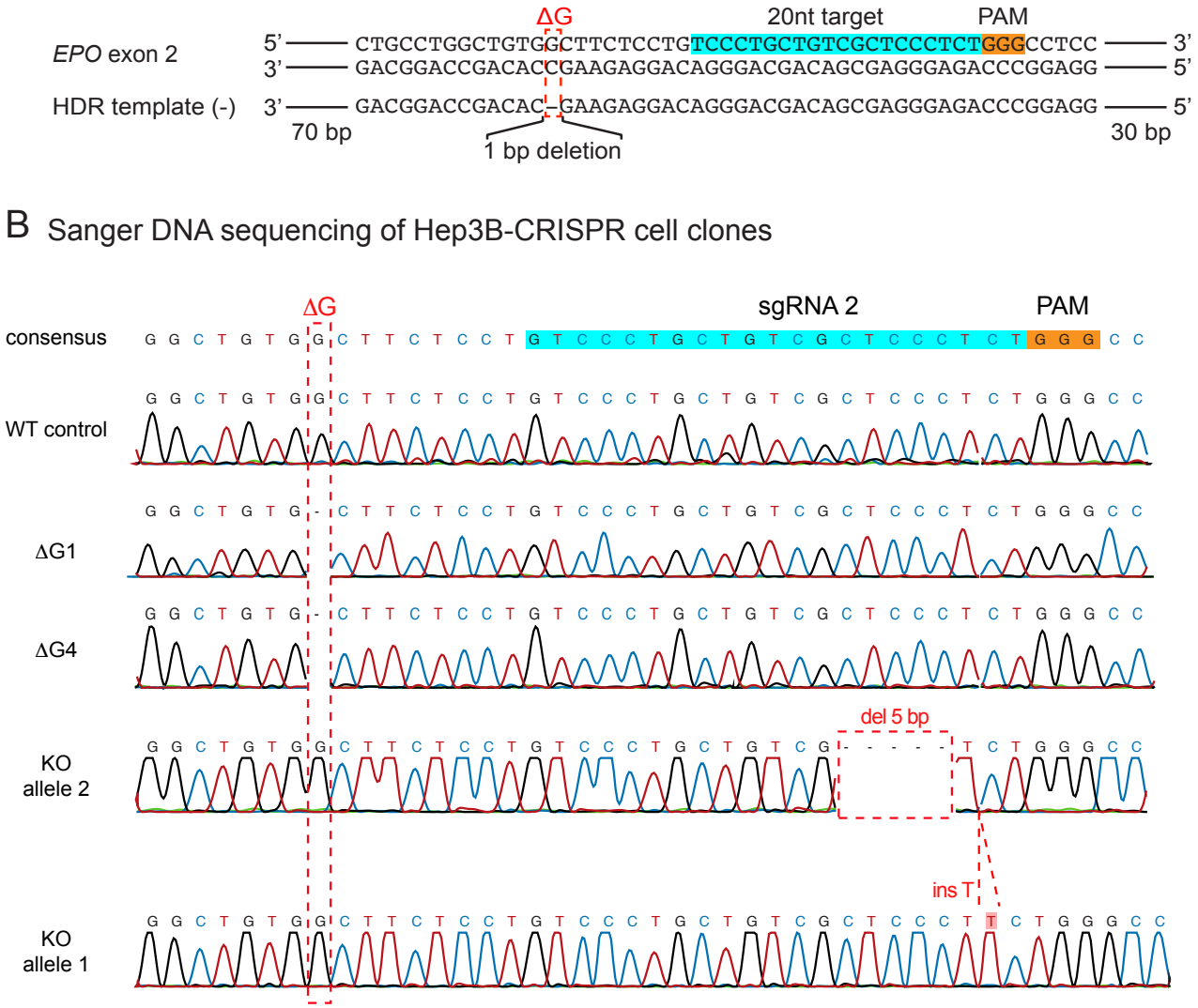
Statistical significance of the data was tested with one-way ANOVA or two-way ANOVA and analysed using GraphPad Prism 7 software with the following significance levels: $p<0.05$= *, $p<0.01$= ** and $p<0.001$= ***.

Supplementary Figure S1



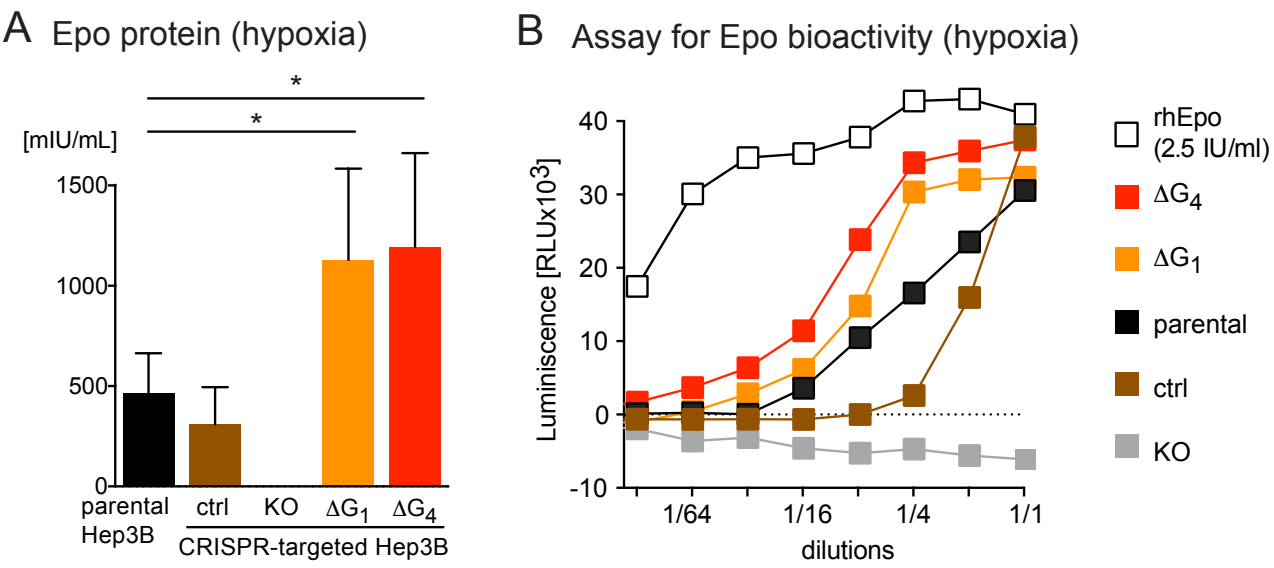
Supplementary Figure S1. Co-segregation of microsatellite markers on chromosome 7q22.1 with the erythrocytosis phenotype. Schematic drawing of the pedigree with unique patient numbers placed under the symbols and *EPO* gene sequence status indicated as wildtype (WT) or mutated (ΔG). Haplotypes for ten polymorphic markers from D7S669 to D7S530 are shown for each family member. The minimal co-segregating region is marked in red and the haplotype linked with the disease is boxed. Physical position of microsatellite markers (STR) on chromosome 7 is shown. Microsatellite markers defining the co-segregating region are shown in red. Two SNPs marking the borders of the co-segregating region identified by SNP array (see Figure 1C) are shown for the comparison.

Supplementary Figure S2



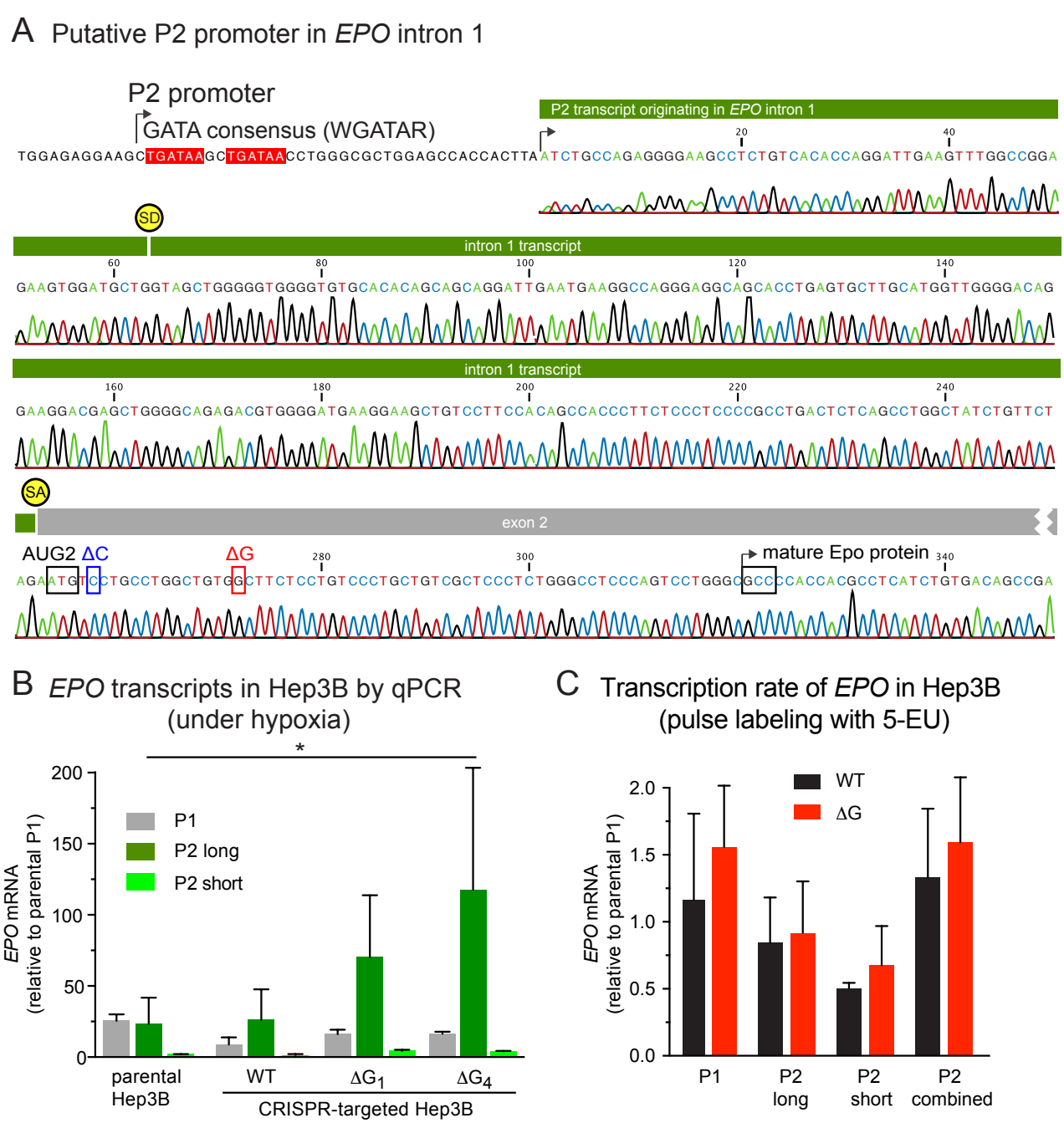
Supplementary Figure S2. **Introducing the EPO ΔG mutation into Hep3B cells by CRISPR.** A) CRISPR-mediated mutagenesis of the *EPO* locus in Hep3B cells. The homology-directed repair (HDR) donor DNA is a 150 base pair (bp) single-stranded oligodeoxynucleotide that is complementary to the target sequence and contains a 1 bp deletion (ΔG), shown in a red box. This modification is flanked by asymmetrical homology arms at 5' and 3' ends. Single-guide RNA targeting sequence is depicted in light blue and protospacer adjacent motif (PAM) in orange. B) Sequencing chromatograms of Hep3B clones modified by CRISPR. Deletions or insertions are marked with red dashed lines. The single-guide RNA (sgRNA) targeting sequence is depicted in light blue and the protospacer adjacent motif (PAM) in orange.

Supplementary Figure S3



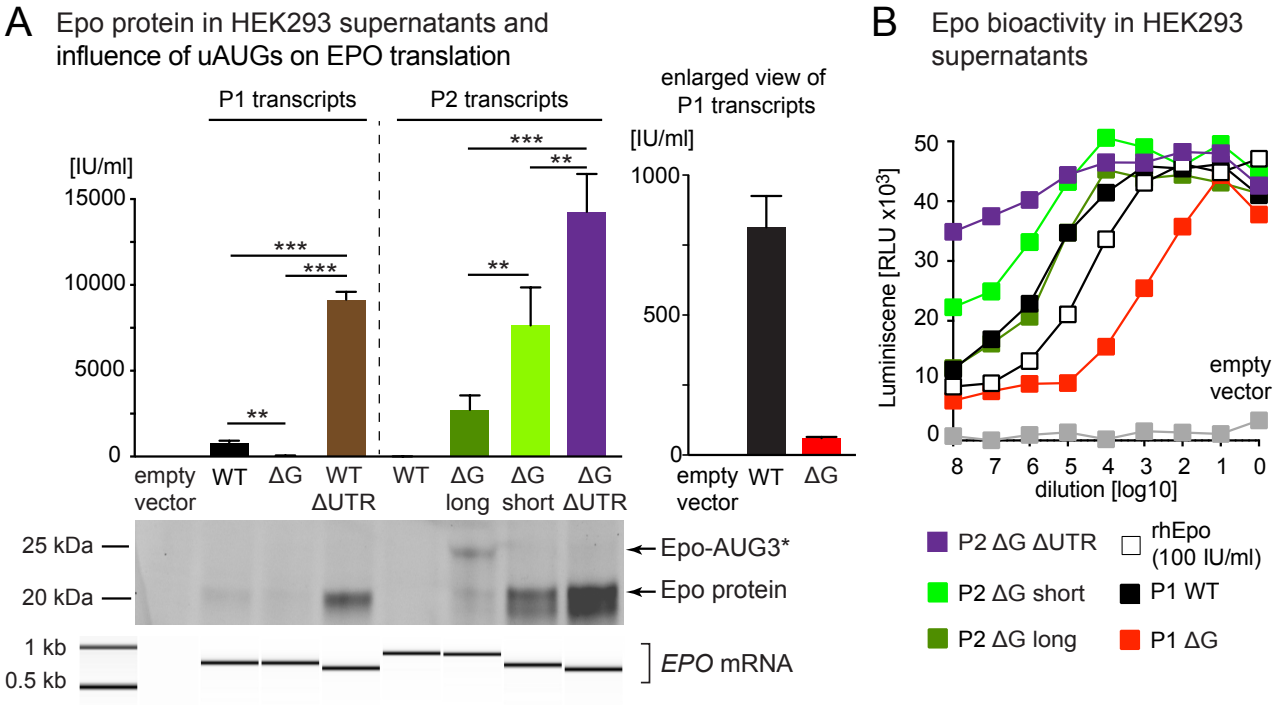
Supplementary Figure S3. **Functional characterization of Hep3B-CRISPR clones under the hypoxia.** A) Epo protein production by Hep3B clones. Supernatants of parental or CRISPR-modified HepB3 cells incubated in hypoxic conditions (1%O_2, 5% CO_2) for 72 hours and were collected and Epo protein was measured by Epo ELISA. The levels of Epo are given as a mean ± SD (n=3), $p < 0.05$ = *, as determined by one-way ANOVA. B) Biological activity of Epo in supernatants of parental or CRISPR-modified HepB3 cells incubated in hypoxic conditions (1%O_2, 5% CO_2) for 72 hours was measured by a proliferation assay with BaF3-EpoR cell line. Serial dilutions of 2.5 IU/ml of recombinant human Epo were used for the reference, n=3.

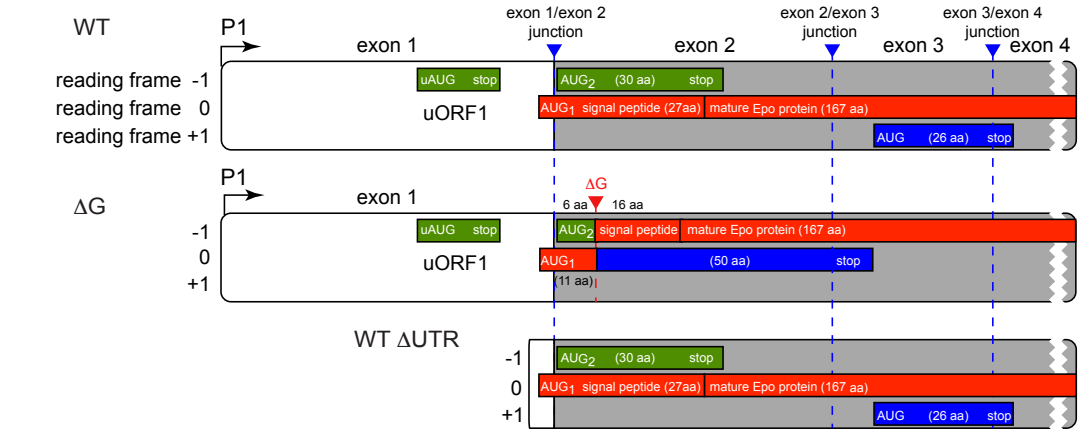
Supplementary Figure S4

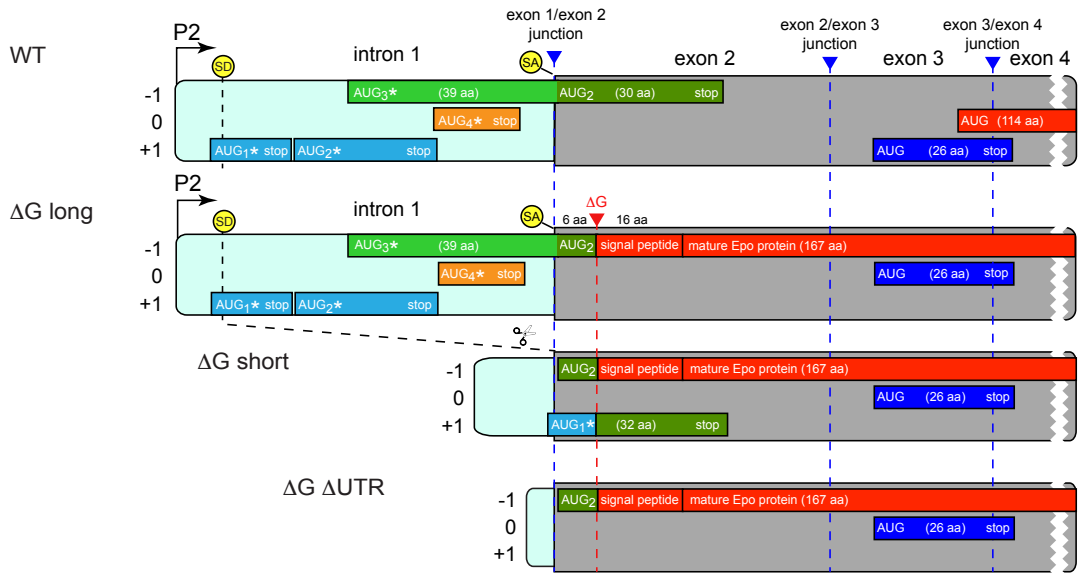


Supplementary Figure S4. **Characterization of EPO P1 and P2 transcripts.** A) Sequences of a putative promoter P2 in *EPO* intron 1 and chromatograms of P2 mRNA sequences identified by 5'-RACE of human liver RNA. The location of a putative P2 promoter is marked by an arrow and two GATA-consensus motifs are highlighted in red. The sequence of the longest P2 transcript identified by 5'-RACE is indicated by green bar, exon 2 sequences by a grey bar. A short version of P2 transcript is produced by splicing out the region of intron 1 marked by splice donor (SD) and splice acceptor (SA) sites (yellow circles). B) Quantification of different *EPO* transcripts in Hep3B cells by qRT-PCR in hypoxic conditions. Total RNA from CRISPR-modified and parental HepB3 cell lines incubated in hypoxic conditions (1%O_2, 5% CO_2) for 72 hours was used. Values are means ± SD (n=2), $p < 0.05$ = *, as determined by two-way ANOVA. C) Transcription rate of *EPO* mRNAs. Quantification of nascent wildtype (WT) and mutant (ΔG) *EPO* mRNAs in Hep3B parental cells (black bars) and ΔG cells (red bars). 5-EU labelled transcripts were quantified by qRT-PCR. Values are means ± SD (n=3). (D) Determination of the stability of *EPO* transcripts. The stability of *EPO* mRNAs was examined in Hep3B parental cells (black points) or ΔG cells (red points). *EPO* expression levels were measured by qRT-PCR and are displayed as a percentage of the levels at the time point 0 h, which is defined as 100%. Values are means ± SD (n=3). Decay curves are non-linear one phase decay fit. Half-life of mRNA ($t_{1/2}$) is an estimation of time, in which 50% of existing mRNA molecules is degraded.

Supplementary Figure S5

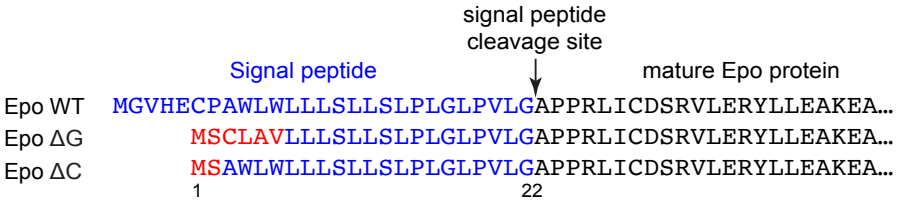




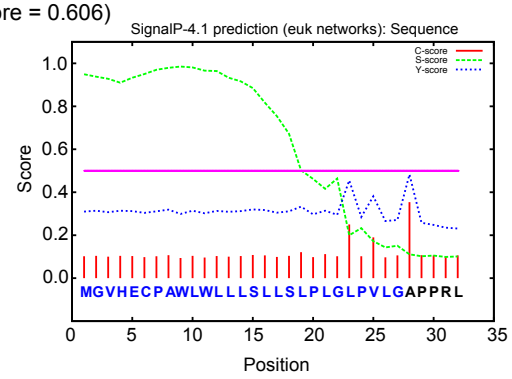


Supplementary Figure S5. Epo protein production by *EPO* P1 and P2 transcripts. A) Upper part: Epo protein measured by immunoassay in supernatants of HEK293 transfected with cDNAs representing P1 or P2 transcripts or empty vector. The levels of Epo are given as a mean ± SD (n=4), $p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = *** as determined by one-way ANOVA. An enlarged view of the results of the P1 transcripts is shown to the right. Lower part: *In vitro* transcription-translation of *EPO* mRNAs. Equal amounts of *in vitro* transcribed *EPO* mRNA variants (bottom) were used for *in vitro* translation in wheat germ lysates in the presence of fluorescently labelled lysine. Epo proteins separated by SDS-gel electrophoresis and visualized by fluorescent imager are shown on top. The predicted molecular weight of wildtype Epo protein including signal peptide is 21.3 kDa. Epo protein with ΔG mutation is shortened by 5 amino acids and expected to be 20.6 kDa. The cause for lower MW bands in the ΔG short and ΔUTR lanes is unknown. A higher molecular weight band of ~25 kDa present in the lane for P2 long ΔG mRNA most likely represents Epo protein initiating from the upstream AUG3* that is in frame with the Epo reading frame. B) Supernatants of transfected HEK293 cells contain biologically active Epo protein. Epo activity was measured by a proliferation assay with BaF3-EpoR cell line. Serial dilutions of 100 IU/ml of recombinant human Epo were used for reference, n=4. C) Schematic drawing showing the translational reading frames of *EPO* P1 and P2 transcripts and the consequences of the ΔG mutation on the *EPO* expression. The upper panel represents the P1 transcript with or without ΔG mutation and an artificial P1 ΔUTR construct lacking uORF1. The lower panel shows the P2 transcripts with or without ΔG mutation and an artificial P2 ΔG ΔUTR construct starting with AUG2 as a first available initiation codon. The translational start codons (AUG) are numbered and placed in the corresponding reading frames. AUGs located in intron 1 are marked by an asterisk (*). Red dotted line denotes the position of the ΔG mutation. Splice donor (SD) and splice acceptor (SA) sites in intron 1 are drawn as yellow circles and the splicing event that generates the EPO short mRNA is marked by dashed lines. Dashed blue lines mark the position of spliced intron-exon 2 junctions.

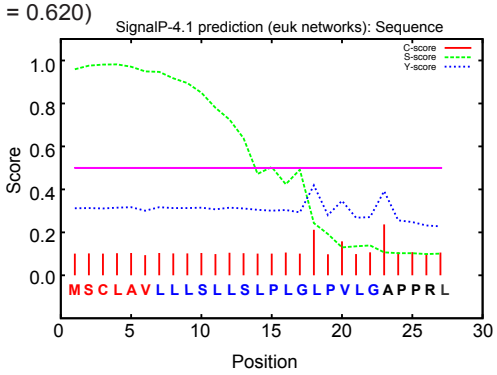
Supplementary Figure S6 Predictions by the SignalP algorithm



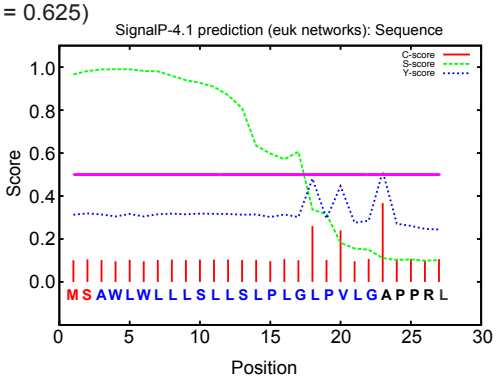
Epo WT (D-score = 0.606)



Epo ΔG (D-score = 0.620)

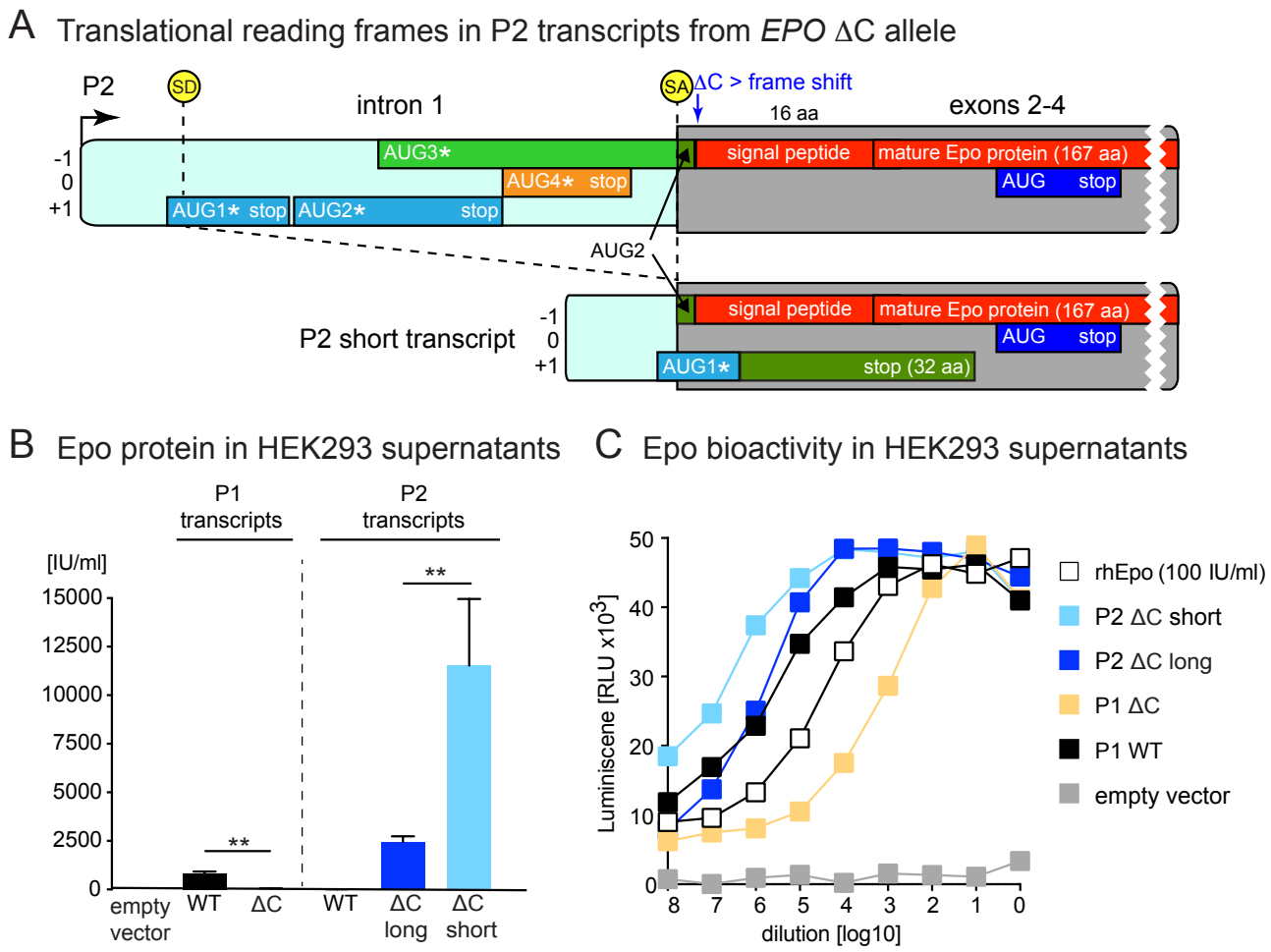


Epo ΔC (D-score = 0.625)



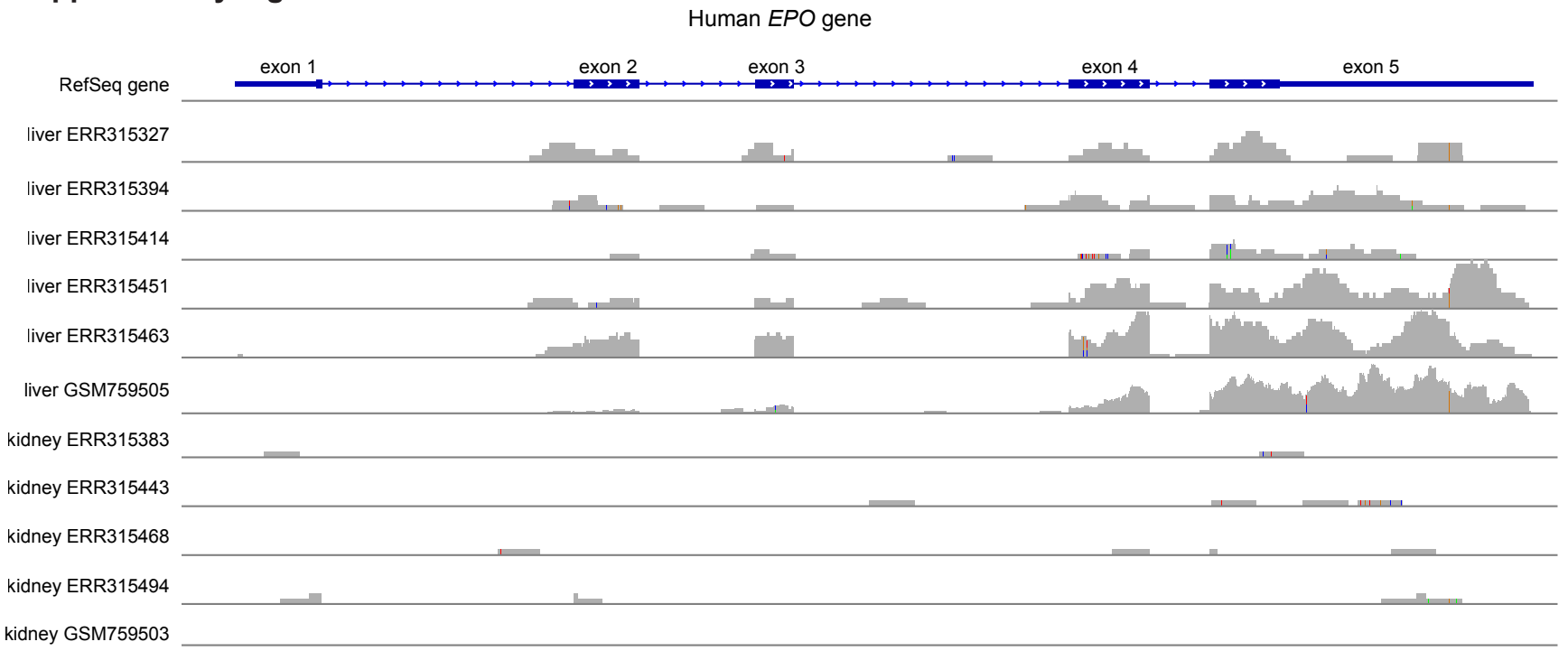
Supplementary Figure S6. Functional prediction of Epo WT, Epo ΔG and Epo ΔC signal peptides. Comparison of Epo WT, Epo ΔG and Epo ΔC signal peptides (SPs) predicted by SignalP 4.1 algorithm. [13] Algorithm produces three output scores for each amino acid in the sequence. The C-score (raw cleavage site score) is high at the position immediately after the cleavage site, S-score (signal peptide score) distinguishes SPs from the mature part of the proteins. Y-score (combined cleavage site score) is a combination of the C-score and the slope of the S-score. A weighted average of the mean S- and maximal Y-scores is used to generate a D-score (discrimination score) that discriminates signal peptides from non-signal peptides. D-score > 0.450 predicts a signal peptide. SPs are marked in blue, mature Epo peptide in black. N-terminal parts of Epo ΔG and Epo ΔC SPs that differ from WT are highlighted in red.

Supplementary Figure S7



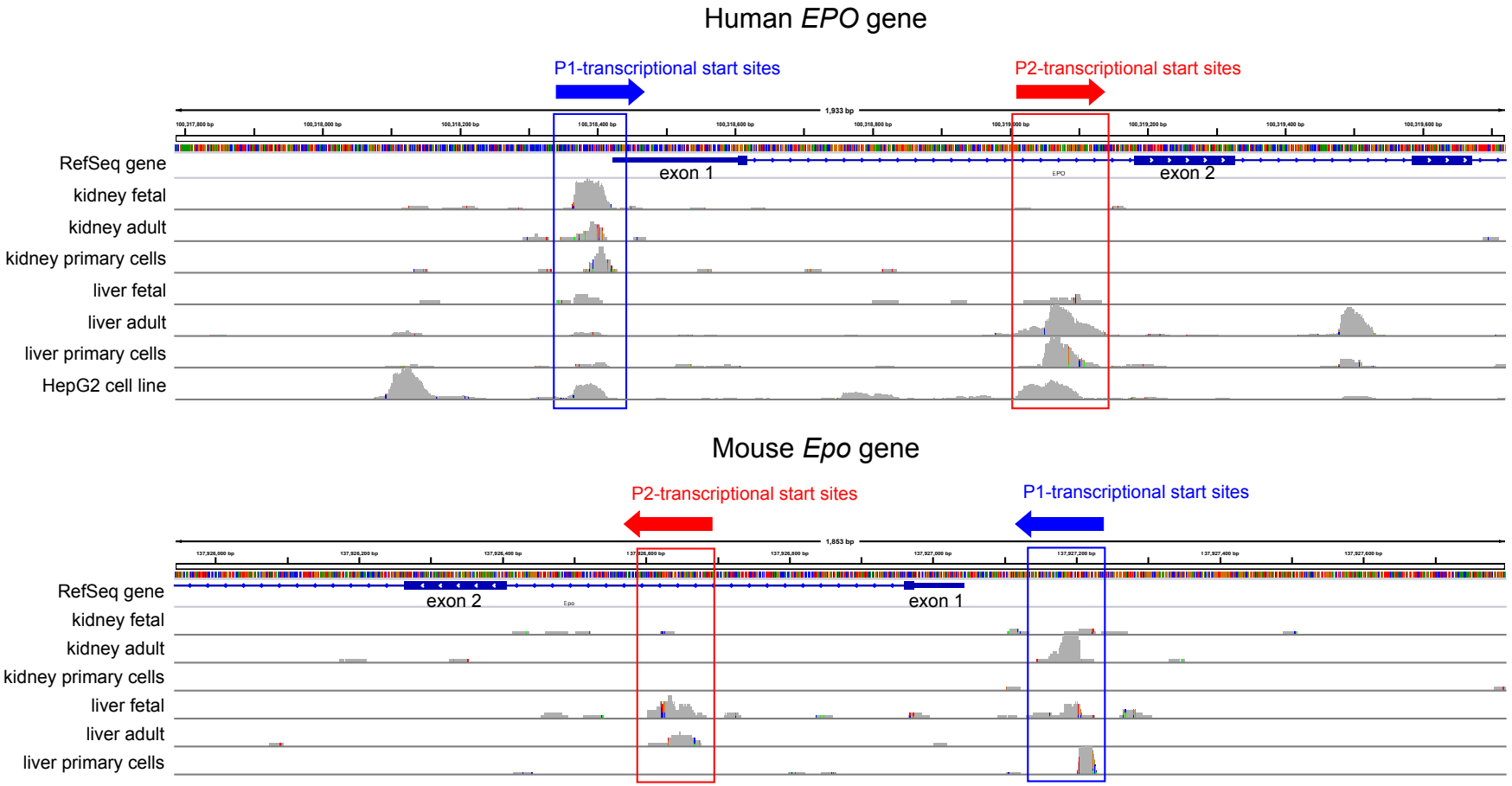
Supplementary Figure S7. *EPO* c.19delC (ΔC) mutation also causes overproduction of Epo. A) Schematic drawing showing the location and the consequences of the ΔC mutation on the *EPO* P2 transcripts. B) Measurement of Epo protein levels and C) Epo biological activity of the supernatants from HEK293 cells transfected with *EPO* P1 and P2 transcripts cDNA with ΔC mutant constructs, similar to Figure S5A and S5B. Note that the results of the wildtype P1 *EPO* construct, the wildtype ΔUTR construct, and the empty vector and rh*EPO* controls are duplicated and shown for reference also in Supplementary Figure S5.

Supplementary Figure S8



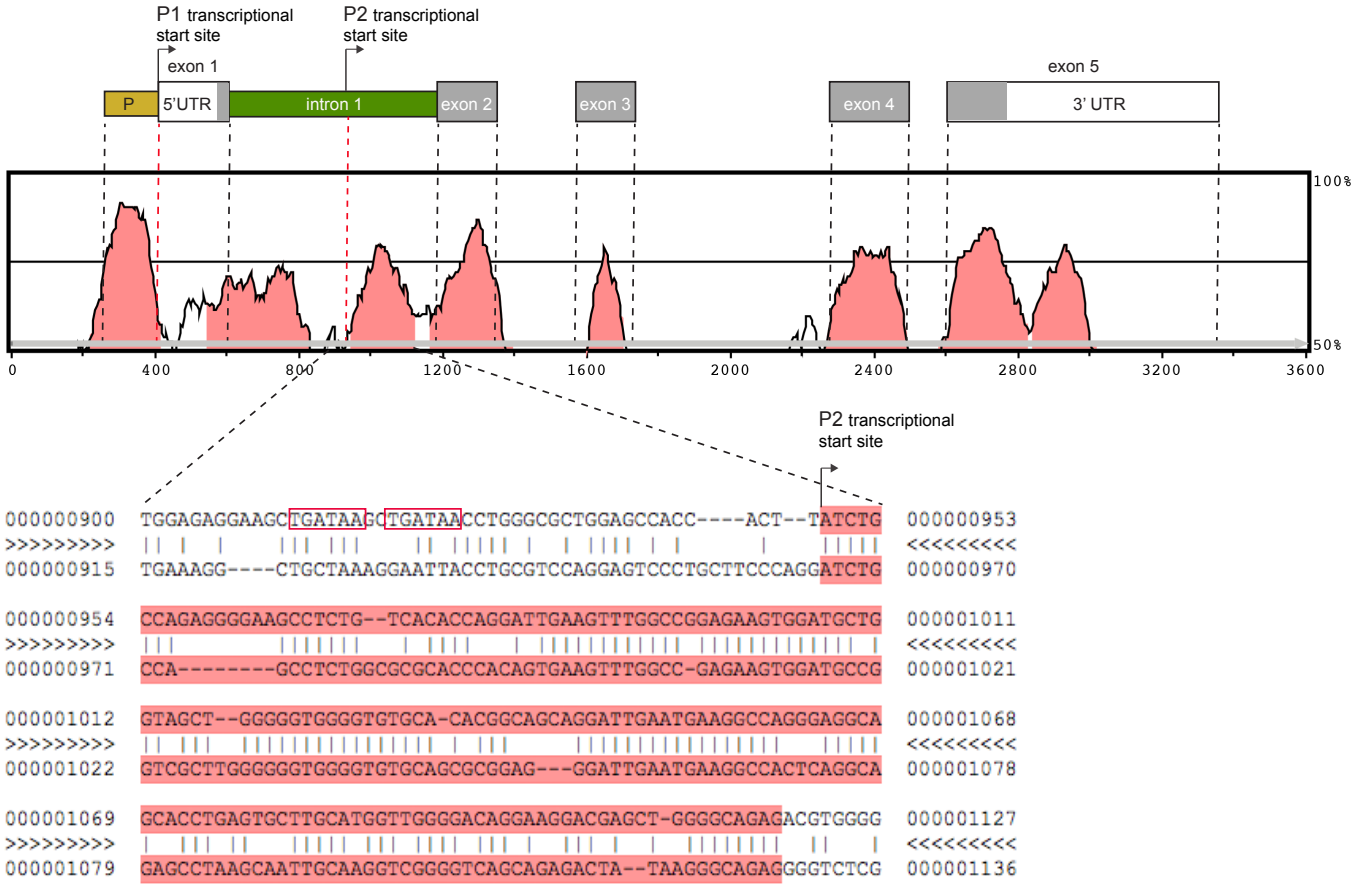
Supplementary Figure S8. ***EPO* P2 transcripts in RNA sequencing data of human kidney and liver.** RNAseq reads were downloaded from the Human Protein Atlas [5] or Illumina Body Map, data were processed as described in Methods section and read coverage was generated with IGV v2.3.[12]

Supplementary Figure S9



Supplementary Figure S9. ***EPO*** **P1 and P2 transcripts in CAGE-sequencing data of human and mouse kidney and liver.** CAGE-sequencing data from FANTOM5 project [10,11] mapped to the reference genomes. Read coverage was generated with IGV v2.3.[12] Transcriptional start sites for P1 are marked with a blue frame and P2 transcripts are marked with a red frame. Arrows denote 5'->3' orientation.

Supplementary Figure S10



Supplementary Figure S10. Intron 1 of *EPO* gene is conserved between the human and the mouse. Alignment of human and mouse *EPO* DNA sequences as calculated by mVISTA [14], using 100 bp window size is shown in upper panel. Regions in red show at least 70% homology. Promoter (P) region is depicted as a yellow box, CDS as grey boxes, 5'- and 3'-untranslated regions as white boxes and intron 1 as a green box. P1 and P2 translational start sites (TSS) are showed with the arrow and red dotted line. Exons are showed with the black dotted line. Lower panel shows the closer look on the part of the intron 1, upper sequence is of human origin, lower of a mouse origin. Two GATA-consensus motifs are shown in red frames, P2 TSS is shown with an arrow and sequence in red is conserved between the human and the mouse.

Supplementary Table 1. Blood parameters of the family members.

UPN	Gender	Age at diagnosis	ΔG	Hb g/L	WBC $x10^9$/L	Plt $x10^9$/L	Epo IU/L
WG01	F	24	ΔG	197	8.2	176	14.6
WG02	M	25	ΔG	204	6.2	238	17.2
WG04	F	16	ΔG	183*	5.6	262	14.5
WG05	M	27	ΔG	167	6.8	197	9.3
WG06	F	7	ΔG	174*	5.9	194	37.7
WG08	F	53	ΔG	198	7.1	176	39.0
WG11	M	21	ΔG	215	3.1	173	51.2
WG14	F	8	ΔG	180*	5.9	386	84.6
WG15	F	15	ΔG	187*	5.2	365	12.6
WG20	F	1.5	ΔG	200	9.6	263	na
WG03	M		WT	155	8.0	176	8.8
WG09	M		WT	158	7.0	205	11.0
WG12	F		WT	152	8.5	336	na
WG13	M		WT	146	9.0	280	na
WG17	M		WT	145	5.8	242	na
WG18	M		WT	152	6.5	205	na
WG19	F		WT	131	5.7	280	13.4
WG21	M		WT	117	8.9	401	7.1

* Hb are values at diagnosis, except for WG04, WG06, WG14 and WG15 that represent the highest value recorded after the start of phlebotomy

Supplementary Table 2. Gene list of co-segregating region identified by SNP array.

Chr	Start position	End position	Gene name	Gene description
chr7	95401817	95727736	DYNC1I1	Homo sapiens dynein, cytoplasmic 1, intermediate chain 1 (DYNC1I1), transcript variant 1, mRNA.
chr7	95749531	95951459	SLC25A13	Homo sapiens solute carrier family 25 (aspartate/glutamate carrier), member 13 (SLC25A13), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
chr7	95848973	95849068	MIR591	Homo sapiens microRNA 591 (MIR591), microRNA.
chr7	95936144	95936164	JB074828	Sequence 777 from Patent WO2010139812.
chr7	95970451	95970558	U6	Rfam model RF00026 hit found at contig region AC004458.1/138309-138203
chr7	96110937	96132835	C7orf76	Homo sapiens chromosome 7 open reading frame 76 (C7orf76), transcript variant 1, mRNA.
chr7	96250968	96293650	LOC100506136	Homo sapiens uncharacterized LOC100506136 (LOC100506136), non-coding RNA.
chr7	96318078	96339203	SHFM1	Homo sapiens split hand/foot malformation (ectrodactyly) type 1 (SHFM1), mRNA.
chr7	96569378	96569676	Metazoa_SRP	Rfam model RF00017 hit found at contig region AC004774.1/46257-45960
chr7	96594613	96594793	Evf1_1	Rfam model RF01887 hit found at contig region AC004774.1/71195-71374
chr7	96594800	96595007	Evf1_2	Rfam model RF01888 hit found at contig region AC004774.1/71382-71588
chr7	96597826	96643377	DLX6-AS1	Homo sapiens DLX6 antisense RNA 1 (DLX6-AS1), non-coding RNA.
chr7	96635289	96640352	DLX6	Homo sapiens distal-less homeobox 6 (DLX6), mRNA.
chr7	96649701	96654143	DLX5	Homo sapiens distal-less homeobox 5 (DLX5), mRNA.
chr7	96745904	96811075	ACN9	Homo sapiens ACN9 homolog (S. cerevisiae) (ACN9), mRNA.
chr7	97228244	97228572	7SK	Rfam model RF00100 hit found at contig region AC003085.2/78764-78437
chr7	97361270	97369784	TAC1	Homo sapiens tachykinin, precursor 1 (TAC1), transcript variant beta, mRNA.
chr7	97481428	97501854	ASNS	Homo sapiens asparagine synthetase (glutamine-hydrolyzing) (ASNS), transcript variant 2, mRNA.
chr7	97503666	97601638	MGC72080	Homo sapiens MGC72080 pseudogene (MGC72080), non-coding RNA.
chr7	97595943	97596475	BC122864	Homo sapiens cDNA clone IMAGE:40108680.
chr7	97614012	97619416	OCM2	Homo sapiens oncomodulin 2 (OCM2), mRNA.
chr7	97736196	97838944	LMTK2	Homo sapiens lemur tyrosine kinase 2 (LMTK2), mRNA.
chr7	97841567	97842271	BHLHA15	Homo sapiens basic helix-loop-helix family, member a15 (BHLHA15), mRNA.
chr7	97844754	97881563	TECPR1	Homo sapiens tectonin beta-propeller repeat containing 1 (TECPR1), mRNA.
chr7	97910978	97920839	BRI3	Homo sapiens brain protein I3 (BRI3), transcript variant 1, mRNA.
chr7	97920961	98030427	BAIAP2L1	Homo sapiens BAI1-associated protein 2-like 1 (BAIAP2L1), mRNA.
chr7	97937196	97938205	hCG_2023280	SubName: Full=HCG2023280; SubName: Full=cDNA FLJ30064 fis, clone ADRGL2000323;
chr7	98246596	98259181	NPTX2	Homo sapiens neuronal pentraxin II (NPTX2), mRNA.
chr7	98444110	98467673	TMEM130	Homo sapiens transmembrane protein 130 (TMEM130), transcript variant 1, mRNA.

chr7	98476112	98610866	TRRAP	Homo sapiens transformation/transcription domain-associated protein (TRRAP), transcript variant 1, mRNA.
chr7	98479272	98479352	MIR3609	Homo sapiens microRNA 3609 (MIR3609), microRNA.
chr7	98625057	98741743	SMURF1	Homo sapiens SMAD specific E3 ubiquitin protein ligase 1 (SMURF1), transcript variant 1, mRNA.
chr7	98771196	98805089	KPNA7	Homo sapiens karyopherin alpha 7 (importin alpha 8) (KPNA7), mRNA.
chr7	98870923	98895594	MYH16	Homo sapiens myosin, heavy chain 16 pseudogene (MYH16), non-coding RNA.
chr7	98923495	98963885	ARPC1A	Homo sapiens actin related protein 2/3 complex, subunit 1A, 41kDa (ARPC1A), transcript variant 1, mRNA.
chr7	98972297	98992404	ARPC1B	Homo sapiens actin related protein 2/3 complex, subunit 1B, 41kDa (ARPC1B), mRNA.
chr7	98992297	99006305	PDAP1	Homo sapiens PDGFA associated protein 1 (PDAP1), mRNA.
chr7	99006600	99017239	BUD31	Homo sapiens BUD31 homolog (S. cerevisiae) (BUD31), mRNA.
chr7	99014361	99063824	ATP5J2-PTCD1	Homo sapiens ATP5J2-PTCD1 readthrough (ATP5J2-PTCD1), mRNA.
chr7	99036562	99054996	CPSF4	Homo sapiens cleavage and polyadenylation specific factor 4, 30kDa (CPSF4), transcript variant 1, mRNA.
chr7	99067306	99067378	TRNA_Trp	transfer RNA Trp (anticodon CCA)
chr7	99070514	99085217	ZNF789	Homo sapiens zinc finger protein 789 (ZNF789), transcript variant 1, mRNA.
chr7	99090853	99097877	ZNF394	Homo sapiens zinc finger protein 394 (ZNF394), mRNA.
chr7	99102272	99131445	ZKSCAN5	Homo sapiens zinc finger with KRAB and SCAN domains 5 (ZKSCAN5), transcript variant 2, mRNA.
chr7	99143922	99149757	FAM200A	Homo sapiens family with sequence similarity 200, member A (FAM200A), mRNA.
chr7	99156447	99174076	ZNF655	Homo sapiens zinc finger protein 655 (ZNF655), transcript variant 8, mRNA.
chr7	99195901	99208456	LOC100289187	Homo sapiens transmembrane protein 225-like (LOC100289187), transcript variant 2, mRNA.
chr7	99214570	99230030	ZSCAN25	Homo sapiens zinc finger and SCAN domain containing 25 (ZSCAN25), mRNA.
chr7	99282301	99332819	CYP3A7-CYP3AP1	Homo sapiens CYP3A7-CYP3AP1 readthrough (CYP3A7-CYP3AP1), mRNA.
chr7	99354582	99381811	CYP3A4	Homo sapiens cytochrome P450, family 3, subfamily A, polypeptide 4 (CYP3A4), transcript variant 1, mRNA.
chr7	99425635	99464173	CYP3A43	Homo sapiens cytochrome P450, family 3, subfamily A, polypeptide 43 (CYP3A43), transcript variant 1, mRNA.
chr7	99473684	99474656	OR2AE1	Homo sapiens olfactory receptor, family 2, subfamily AE, member 1 (OR2AE1), mRNA.
chr7	99488029	99517223	TRIM4	Homo sapiens tripartite motif containing 4 (TRIM4), transcript variant beta, mRNA.
chr7	99520891	99527243	GJC3	Homo sapiens gap junction protein, gamma 3, 30.2kDa (GJC3), mRNA.
chr7	99564349	99573735	AZGP1	Homo sapiens alpha-2-glycoprotein 1, zinc-binding (AZGP1), mRNA.
chr7	99578384	99591588	AZGP1P1	Homo sapiens alpha-2-glycoprotein 1, zinc-binding pseudogene 1 (AZGP1P1), non-coding RNA.
chr7	99613218	99635403	ZKSCAN1	Homo sapiens zinc finger with KRAB and SCAN domains 1 (ZKSCAN1), mRNA.
chr7	99647416	99662663	ZSCAN21	Homo sapiens zinc finger and SCAN domain containing 21 (ZSCAN21), mRNA.
chr7	99667593	99679387	ZNF3	Homo sapiens zinc finger protein 3 (ZNF3), transcript variant 2, mRNA.
chr7	99686582	99689822	COPS6	Homo sapiens COP9 signalosome subunit 6 (COPS6), mRNA.
chr7	99690403	99699427	MCM7	Homo sapiens minichromosome maintenance complex component 7 (MCM7), transcript variant 1, mRNA.
chr7	99691182	99691266	MIR25	Homo sapiens microRNA 25 (MIR25), microRNA.

chr7	99691390	99691470	MIR93	Homo sapiens microRNA 93 (MIR93), microRNA.
chr7	99691437	99691459	DD413568	Sequence 150 from Patent WO2004076622.
chr7	99691615	99691697	MIR106B	Homo sapiens microRNA 106b (MIR106B), microRNA.
chr7	99699129	99704803	AP4M1	Homo sapiens adaptor-related protein complex 4, mu 1 subunit (AP4M1), mRNA.
chr7	99704692	99717481	TAF6	Homo sapiens TAF6 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 80kDa (TAF6), transcript variant 5, mRNA.
chr7	99717264	99723128	CNPY4	Homo sapiens canopy 4 homolog (zebrafish) (CNPY4), mRNA.
chr7	99724319	99726121	MBLAC1	Homo sapiens metallo-beta-lactamase domain containing 1 (MBLAC1), mRNA.
chr7	99746529	99751833	LAMTOR4	Homo sapiens late endosomal/lysosomal adaptor, MAPK and MTOR activator 4 (LAMTOR4), mRNA.
chr7	99752042	99756302	C7orf43	Homo sapiens chromosome 7 open reading frame 43 (C7orf43), mRNA.
chr7	99756864	99766373	GAL3ST4	Homo sapiens galactose-3-O-sulfotransferase 4 (GAL3ST4), mRNA.
chr7	99767228	99774990	GPC2	Homo sapiens glypican 2 (GPC2), mRNA.
chr7	99775537	99812010	STAG3	Homo sapiens stromal antigen 3 (STAG3), mRNA.
chr7	99798277	99869855	GATS	Homo sapiens GATS, stromal antigen 3 opposite strand (GATS), transcript variant 1, mRNA.
chr7	99816871	99819111	PVRIG	Homo sapiens poliovirus receptor related immunoglobulin domain containing (PVRIG), mRNA.
chr7	99905324	99919819	SPDYE3	Homo sapiens speedy/RINGO cell cycle regulator family member E3 (SPDYE3), mRNA.
chr7	99917162	99917190	DQ601342	Homo sapiens piRNA piR-59725, complete sequence.
chr7	99918262	99933930	PMS2P1	Homo sapiens postmeiotic segregation increased 2 pseudogene 1 (PMS2P1), non-coding RNA.
chr7	99927418	99930239	PMS2P1	Homo sapiens postmeiotic segregation increased 2 pseudogene 1 (PMS2P1), non-coding RNA.
chr7	99933766	99949523	PILRB	Homo sapiens paired immunoglobin-like type 2 receptor beta (PILRB), transcript variant 2, non-coding RNA.
chr7	99955625	99965454	PILRB	Homo sapiens paired immunoglobin-like type 2 receptor beta (PILRB), transcript variant 3, mRNA.
chr7	99971067	99997722	PILRA	Homo sapiens paired immunoglobin-like type 2 receptor alpha (PILRA), transcript variant 1, mRNA.
chr7	99998494	100026431	ZCWPW1	Homo sapiens zinc finger, CW type with PWWP domain 1 (ZCWPW1), transcript variant 1, mRNA.
chr7	100027253	100031749	MEPCE	Homo sapiens methylphosphate capping enzyme (MEPCE), transcript variant 1, mRNA.
chr7	100032911	100034094	PPP1R35	Homo sapiens protein phosphatase 1, regulatory subunit 35 (PPP1R35), mRNA.
chr7	100054237	100061894	C7orf61	Homo sapiens chromosome 7 open reading frame 61 (C7orf61), mRNA.
chr7	100064141	100076902	TSC22D4	Homo sapiens TSC22 domain family, member 4 (TSC22D4), mRNA.
chr7	100081549	100092424	NYAP1	Homo sapiens neuronal tyrosine-phosphorylated phosphoinositide-3-kinase adaptor 1 (NYAP1), mRNA.
chr7	100136833	100165843	AGFG2	Homo sapiens ArfGAP with FG repeats 2 (AGFG2), mRNA.
chr7	100169852	100171270	SAP25	Homo sapiens Sin3A-associated protein, 25kDa (SAP25), mRNA.
chr7	100171633	100183776	LRCH4	Homo sapiens leucine-rich repeats and calponin homology (CH) domain containing 4 (LRCH4), mRNA.
chr7	100187023	100201661	PCOLCE-AS1	Homo sapiens PCOLCE antisense RNA 1 (PCOLCE-AS1), non-coding RNA.
chr7	100187195	100198740	FBXO24	Homo sapiens F-box protein 24 (FBXO24), transcript variant 3, mRNA.
chr7	100199881	100205798	PCOLCE	Homo sapiens procollagen C-endopeptidase enhancer (PCOLCE), mRNA.
chr7	100210113	100213000	MOSPD3	Homo sapiens motile sperm domain containing 3 (MOSPD3), transcript variant 1, mRNA.

chr7	100218038	100239201	TFR2	Homo sapiens transferrin receptor 2 (TFR2), transcript variant 1, mRNA.
chr7	100240725	100254084	ACTL6B	Homo sapiens actin-like 6B (ACTL6B), mRNA.
chr7	100254183	100264332	AK055267	Homo sapiens cDNA FLJ30705 fis, clone FCBBF2001116.
chr7	100271362	100276792	GNB2	Homo sapiens guanine nucleotide binding protein (G protein), beta polypeptide 2 (GNB2), mRNA.
chr7	100277129	100286870	GIGYF1	Homo sapiens GRB10 interacting GYF protein 1 (GIGYF1), mRNA.
chr7	100303675	100305123	POP7	Homo sapiens processing of precursor 7, ribonuclease P/MRP subunit (S. cerevisiae) (POP7), mRNA.
chr7	100318422	100321323	EPO	Homo sapiens erythropoietin (EPO), mRNA.
chr7	100331248	100395419	ZAN	Homo sapiens zonadhesin (ZAN), transcript variant 6, mRNA.
chr7	100400186	100425143	EPHB4	Homo sapiens EPH receptor B4 (EPHB4), mRNA.
chr7	100450340	100464634	SLC12A9	Homo sapiens solute carrier family 12 (potassium/chloride transporters), member 9 (SLC12A9), transcript variant 1, mRNA.
chr7	100464949	100471076	TRIP6	Homo sapiens thyroid hormone receptor interactor 6 (TRIP6), mRNA.
chr7	100472700	100486285	SRRT	Homo sapiens serrate RNA effector molecule homolog (Arabidopsis) (SRRT), transcript variant 1, mRNA.
chr7	100486343	100487339	UFSP1	Homo sapiens UFM1-specific peptidase 1 (non-functional) (UFSP1), mRNA.
chr7	100487614	100493541	ACHE	Homo sapiens acetylcholinesterase (ACHE), transcript variant E4-E5, mRNA.
chr7	100488709	100488760	DJ051769	Sequence 22 from Patent WO0136627.
chr7	100550749	100611111	MUC3A	SubName: Full=Mucin-3A; Flags: Fragment;
chr7	100606877	100611410	AK096803	Homo sapiens cDNA FLJ39484 fis, clone PROST2014925.
chr7	100612903	100662230	MUC12	Homo sapiens mucin 12, cell surface associated (MUC12), mRNA.
chr7	100663363	100702140	MUC17	Homo sapiens mucin 17, cell surface associated (MUC17), mRNA.
chr7	100728785	100733889	TRIM56	Homo sapiens tripartite motif containing 56 (TRIM56), mRNA.
chr7	100770369	100782547	SERPINE1	Homo sapiens serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1 (SERPINE1), mRNA.
chr7	100797685	100804557	AP1S1	Homo sapiens adaptor-related protein complex 1, sigma 1 subunit (AP1S1), mRNA.
chr7	100802753	100802836	MIR4653	Homo sapiens microRNA 4653 (MIR4653), microRNA.
chr7	100805789	100808852	VGF	Homo sapiens VGF nerve growth factor inducible (VGF), mRNA.
chr7	100813773	100823557	NAT16	Homo sapiens N-acetyltransferase 16 (GCN5-related, putative) (NAT16), mRNA.
chr7	100839012	100844302	MOGAT3	Homo sapiens monoacylglycerol O-acyltransferase 3 (MOGAT3), mRNA.
chr7	100849257	100861011	PLOD3	Homo sapiens procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3 (PLOD3), mRNA.
chr7	100860984	100867471	ZNHIT1	Homo sapiens zinc finger, HIT-type containing 1 (ZNHIT1), mRNA.
chr7	100875372	100881227	CLDN15	Homo sapiens claudin 15 (CLDN15), transcript variant 2, mRNA.
chr7	100882892	100888371	FIS1	Homo sapiens fission 1 (mitochondrial outer membrane) homolog (S. cerevisiae) (FIS1), nuclear gene encoding mitochondrial protein, mRNA.
chr7	100888045	100888132	Mir_324	Rfam model RF01063 hit found at contig region AC006329.5/14125-14039
chr7	100893615	100895601	AK091729	Homo sapiens cDNA FLJ34410 fis, clone HEART2002129.
chr7	100951620	100954265	BC032716	Homo sapiens cDNA clone IMAGE:5518697, **** WARNING: chimeric clone ****.

chr7	100956647	100965093	RABL5	Homo sapiens RAB, member RAS oncogene family-like 5 (RABL5), transcript variant 1, mRNA.
chr7	100958773	100965093	RABL5	Homo sapiens RAB, member RAS oncogene family-like 5 (RABL5), transcript variant 2, mRNA.
chr7	101006121	101202304	COL26A1	Homo sapiens collagen, type XXVI, alpha 1 (COL26A1), mRNA.
chr7	101208596	101212286	BC027906	Homo sapiens clone 25248 mRNA sequence.
chr7	101256604	101272576	MYL10	Homo sapiens myosin, light chain 10, regulatory (MYL10), mRNA.
chr7	101460881	101901513	CUX1	Homo sapiens cut-like homeobox 1 (CUX1), transcript variant 1, mRNA.
chr7	101928404	101962178	SH2B2	Homo sapiens SH2B adaptor protein 2 (SH2B2), mRNA.
chr7	101936368	101936453	MIR4285	Homo sapiens microRNA 4285 (MIR4285), microRNA.
chr7	101979403	101979439	DQ595418	Homo sapiens piRNA piR-45129, complete sequence.
chr7	101986191	101996889	SPDYE6	Homo sapiens speedy/RINGO cell cycle regulator family member E6 (SPDYE6), mRNA.
chr7	101989114	101989142	DQ601342	Homo sapiens piRNA piR-59725, complete sequence.
chr7	102036803	102067129	PRKRIP1	Homo sapiens PRKR interacting protein 1 (IL11 inducible) (PRKRIP1), mRNA.
chr7	102063377	102067120	AK097289	Homo sapiens cDNA FLJ39970 fis, clone SPLEN2027995.
chr7	102073976	102097268	ORAI2	Homo sapiens ORAI calcium release-activated calcium modulator 2 (ORAI2), transcript variant 2, mRNA.
chr7	102096666	102105321	ALKBH4	Homo sapiens alkB, alkylation repair homolog 4 (E. coli) (ALKBH4), mRNA.
chr7	102105389	102113612	LRWD1	Homo sapiens leucine-rich repeats and WD repeat domain containing 1 (LRWD1), mRNA.
chr7	102106188	102106273	MIR5090	Homo sapiens microRNA 5090 (MIR5090), microRNA.
chr7	102111915	102111978	MIR4467	Homo sapiens microRNA 4467 (MIR4467), microRNA.
chr7	102113547	102119381	POLR2J	Homo sapiens polymerase (RNA) II (DNA directed) polypeptide J, 13.3kDa (POLR2J), mRNA.
chr7	102123585	102158224	RASA4B	Homo sapiens RAS p21 protein activator 4B (RASA4B), mRNA.
chr7	102133833	102135811	AK311374	Homo sapiens cDNA, FLJ18416.
chr7	102178365	102184517	POLR2J3	Homo sapiens polymerase (RNA) II (DNA directed) polypeptide J3 (POLR2J3), mRNA.
chr7	102178365	102213068	POLR2J3	Homo sapiens polymerase (RNA) II (DNA directed) polypeptide J3 (POLR2J3), mRNA.
chr7	102188513	102189359	BC041025	Homo sapiens cDNA clone IMAGE:3959993.
chr7	102191678	102202757	SPDYE2B	Homo sapiens speedy/RINGO cell cycle regulator family member E2B (SPDYE2B), mRNA.
chr7	102201021	102201049	DQ601342	Homo sapiens piRNA piR-59725, complete sequence.
chr7	102220092	102257205	RASA4	Homo sapiens RAS p21 protein activator 4 (RASA4), transcript variant 1, mRNA.
chr7	102233002	102234980	AK311374	Homo sapiens cDNA, FLJ18416.
chr7	102277194	102283238	POLR2J2	Homo sapiens polymerase (RNA) II (DNA directed) polypeptide J2 (POLR2J2), mRNA.
chr7	102277194	102312176	POLR2J2	Homo sapiens polymerase (RNA) II (DNA directed) polypeptide J2 (POLR2J2), mRNA.
chr7	102287607	102288458	BC041025	Homo sapiens cDNA clone IMAGE:3959993.
chr7	102290771	102301847	SPDYE2B	Homo sapiens speedy/RINGO cell cycle regulator family member E2B (SPDYE2B), mRNA.
chr7	102300111	102300139	DQ601342	Homo sapiens piRNA piR-59725, complete sequence.
chr7	102321131	102328837	AK301666	SubName: Full=cDNA FLJ50902, moderately similar to Ras GTPase-activating protein 4;

chr7	102389398	102449672	FAM185A	Homo sapiens family with sequence similarity 185, member A (FAM185A), transcript variant 1, mRNA.
chr7	102453307	102715015	FBXL13	Homo sapiens F-box and leucine-rich repeat protein 13 (FBXL13), transcript variant 1, mRNA.
chr7	102553343	102585556	LRRC17	Homo sapiens leucine rich repeat containing 17 (LRRC17), transcript variant 1, mRNA.
chr7	102715327	102740210	ARMC10	Homo sapiens armadillo repeat containing 10 (ARMC10), transcript variant A, mRNA.
chr7	102740022	102789569	NAPEPLD	Homo sapiens N-acyl phosphatidylethanolamine phospholipase D (NAPEPLD), transcript variant 1, mRNA.
chr7	102781716	102782850	RPL19P12	Homo sapiens ribosomal protein L19 pseudogene 12 (RPL19P12), non-coding RNA.
chr7	102815461	102920759	DPY19L2P2	Homo sapiens dpy-19-like 2 pseudogene 2 (C. elegans) (DPY19L2P2), transcript variant 2, non-coding RNA.
chr7	102913538	102920913	DPY19L2P2	Homo sapiens dpy-19-like 2 pseudogene 2 (C. elegans) (DPY19L2P2), transcript variant 1, non-coding RNA.
chr7	102937872	102955133	PMPCB	Homo sapiens peptidase (mitochondrial processing) beta (PMPCB), nuclear gene encoding mitochondrial protein, mRNA.
chr7	102952920	102985320	DNAJC2	Homo sapiens DnaJ (Hsp40) homolog, subfamily C, member 2 (DNAJC2), transcript variant 1, mRNA.
chr7	102987970	103009842	PSMC2	Homo sapiens proteasome (prosome, macropain) 26S subunit, ATPase, 2 (PSMC2), transcript variant 1, mRNA.
chr7	103014654	103086624	SLC26A5	Homo sapiens solute carrier family 26, member 5 (prestin) (SLC26A5), transcript variant a, mRNA.
chr7	103075440	103075539	Y_RNA	Rfam model RF00019 hit found at contig region AC005064.3/33168-33070
chr7	103112230	103629963	RELN	Homo sapiens reelin (RELN), transcript variant 2, mRNA.
chr7	103124654	103124986	7SK	Rfam model RF00100 hit found at contig region AC005064.3/82615-82284
chr7	103766787	103848495	ORC5	Homo sapiens origin recognition complex, subunit 5 (ORC5), transcript variant 1, mRNA.
chr7	103969103	104549003	LHFPL3	Homo sapiens lipoma HMGIC fusion partner-like 3 (LHFPL3), mRNA.
chr7	103969740	103969760	JB175200	Oligonucleotides for modulating target RNA activity.
chr7	104439446	104444539	LHFPL3-AS1	Homo sapiens LHFPL3 antisense RNA 1 (LHFPL3-AS1), transcript variant 1, non-coding RNA.
chr7	104535074	104567092	LHFPL3-AS2	Homo sapiens LHFPL3 antisense RNA 2 (LHFPL3-AS2), non-coding RNA.
chr7	104581652	104602781	BC061919	Homo sapiens cDNA clone IMAGE:4358995, partial cds.
chr7	104622193	104631612	LOC100216546	Homo sapiens uncharacterized LOC100216546 (LOC100216546), non-coding RNA.
chr7	104650988	104654588	LOC100216545	Homo sapiens uncharacterized LOC100216545 (LOC100216545), non-coding RNA.
chr7	104654636	104754532	KMT2E	Homo sapiens lysine (K)-specific methyltransferase 2E (KMT2E), transcript variant 1, mRNA.
chr7	104748533	104749020	AF520793	Homo sapiens chromosome 7 unknown mRNA.
chr7	104754024	104754437	AF520792	Homo sapiens clone IMAGE: 252897 mRNA sequence.
chr7	104754466	104755465	AF520792	Homo sapiens clone IMAGE: 252897 mRNA sequence.
chr7	104756820	105029377	SRPK2	Homo sapiens SRSF protein kinase 2 (SRPK2), transcript variant 1, mRNA.
chr7	105096959	105162685	PUS7	Homo sapiens pseudouridylate synthase 7 homolog (S. cerevisiae) (PUS7), mRNA.
chr7	105172531	105208124	RINT1	Homo sapiens RAD50 interactor 1 (RINT1), mRNA.
chr7	105206771	105221976	EFCAB10	Homo sapiens EF-hand calcium binding domain 10 (EFCAB10), non-coding RNA.
chr7	105208141	105208836	BC007100	Homo sapiens cDNA clone IMAGE:4279216, **** WARNING: chimeric clone ****.
chr7	105245220	105517031	ATXN7L1	Homo sapiens ataxin 7-like 1 (ATXN7L1), transcript variant 1, mRNA.
chr7	105603656	105676877	CDHR3	Homo sapiens cadherin-related family member 3 (CDHR3), mRNA.

chr7	105730813	105753093	SYPL1	Homo sapiens synaptophysin-like 1 (SYPL1), transcript variant 1, mRNA.
chr7	105848612	105848713	U6	Rfam model RF00026 hit found at contig region AC007032.2/13823-13923
chr7	105888731	105925638	NAMPT	Homo sapiens nicotinamide phosphoribosyltransferase (NAMPT), mRNA.
chr7	106130171	106410653	AF086203	Homo sapiens full length insert cDNA clone ZC44D09.
chr7	106297210	106301634	CCDC71L	Homo sapiens coiled-coil domain containing 71-like (CCDC71L), mRNA.
chr7	106505923	106547592	PIK3CG	Homo sapiens phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit gamma (PIK3CG), mRNA.
chr7	106685177	106802256	PRKAR2B	Homo sapiens protein kinase, cAMP-dependent, regulatory, type II, beta (PRKAR2B), mRNA.
chr7	106809405	106842974	HBP1	Homo sapiens HMG-box transcription factor 1 (HBP1), transcript variant 1, mRNA.
chr7	106842188	107204959	COG5	Homo sapiens component of oligomeric golgi complex 5 (COG5), transcript variant 1, mRNA.
chr7	107110501	107116125	GPR22	Homo sapiens G protein-coupled receptor 22 (GPR22), mRNA.
chr7	107204401	107218968	DUS4L	Homo sapiens dihydrouridine synthase 4-like (S. cerevisiae) (DUS4L), transcript variant 2, mRNA.
chr7	107221203	107263762	BCAP29	Homo sapiens B-cell receptor-associated protein 29 (BCAP29), transcript variant 1, mRNA.
chr7	107243807	107243952	SnoU109	Rfam model RF01233 hit found at contig region AC004839.1/36528-36672
chr7	107296960	107302243	SLC26A4-AS1	Homo sapiens SLC26A4 antisense RNA 1 (SLC26A4-AS1), non-coding RNA.
chr7	107301079	107358252	SLC26A4	Homo sapiens solute carrier family 26, member 4 (SLC26A4), mRNA.
chr7	107312769	107312855	Mir_548	Rfam model RF01061 hit found at contig region AC078937.1/49798-49713
chr7	107384278	107402083	CBLL1	Homo sapiens Cbl proto-oncogene, E3 ubiquitin protein ligase-like 1 (CBLL1), transcript variant 1, mRNA.
chr7	107405911	107443678	SLC26A3	Homo sapiens solute carrier family 26, member 3 (SLC26A3), mRNA.

Supplementary Table 3. Translational start predictions.

	Position	Score	Prediction	Kozak consensus **gccRccAUGG**
P1 transcript				
uAUG	115	0.549	yes	cccGggAUGA
AUG1	182	0.729	yes	gcgGagAUGG
AUG2	196	0.644	yes	gcaCgaAUGT
P2 long transcript				
AUG1*	58	0.263	no	aagTggAUGC
AUG2*	103	0.416	no	gatTgaAUGA
AUG3*	137	0.482	no	gctTgcAUGG
AUG4*	180	0.351	no	gtgGggAUGA
AUG2	254	0.656	yes	gcaCgaAUGT
P2 short transcript				
AUG1*	58	0.314	no	aagTggAUGC
AUG2	65	0.512	yes	gcaCgaAUGT

NetStart 1.0 prediction server [15] has been used to calculate the robustness of different AUGs to initiate the translation in the P1, P2 and P2 short transcripts. Score greater than 0.5 represent a probable translation start.

Supplementary Table 4. Primer list.

***EPO* exon 2 sequencing**	**primer sequence 5'->3'**
EPO exon 2 fwd	GCCACCACTTATCTGCCA
EPO exon 2 rev	AACTCCACCCCAAACCAA

Microsatellite mapping	**primer sequence 5'->3'**
D7S669 (q21.11)	FAM-ATGCAACCTACCCTCAAATG
	gtttcTACGGGTTACCCACATTGCTAT
D7S630	FAM-TCCATTCTGAGGTTTGATGT
	gtttcCCATGGTCTTTTCAATGAAC
D7S657	FAM-GGAATCTGGTAGACTGGTTT
	gtttcCCCTGCCTCTAAAATTATAC
D7S1812	FAM-GGAGGAAGGTGGATAATAATAGG
	gtttcAATGTCACAAGGGTGCAGAT
D7S554	FAM-GTCTAATTACCCACATTTCCCT
	gtttcGTTCCATATTTAAAAGACTCAGTGA
D7S666	FAM-GCCTTCTCAAGCAAATTGAT
	gtttcCCTGATATGTGAGGTAATGAAAGAG
D7S2504	FAM-TGTGGTACAATTTCAGACACATAA
	gtttcCTGGAAACCAGTGTTTTCACTT
D7S2459	FAM-AAGAAGTGCATTGAGACTCC
	gtttcCCGCCTTAGTAAAACCC
D7S486	FAM-AAAGGCCAATGGTATATCCC
	gtttcGCCCAGGTGATTGATAGTGC
D7S530 (q32.2)	FAM-TGCATTTTAGTGGAGCACAG
	gtttcCAGGCATTGGGAACTTTG

CRISPR sgRNA cloning	**primer sequence 5'->3'**
sgRNA #2	CACCGTCCCTGCTGTCGCTCCCTCT
TCCCTGCTGTCGCTCCCTCT	AAACAGAGGGAGCGACAGCAGGGAC

SNP TaqMan genotyping primers	**primer sequence 5'->3'**
hEPO ΔG SNP genotyping fwd	TATCTGTTCTAGAATGTCCTGCCT
hEPO ΔG SNP genotyping rev	GGAGCGACAGCAGGGA
TaqMan probe WT (VIC-NFQ)	CAGGAGAAGCCACAGCC
TaqMan probe MUT (FAM-NFQ)	CAGGAGAAGCACAGCC

GSP primers for 5'RACE	**primer sequence 5'->3'**
GSP1	GATTACGCCAAGCTT**AGCAGTGATTGTTCGGAGTGGAGCAG**
GSP2	GATTACGCCAAGCTT**AGAGCCCGAAGCAGAGTGGTGAGGC**

***EPO* transcripts quantitation**	**primer sequence 5'->3'**
EPO P1 fwd	CGGCCAGGCGCGGAGATG
EPO P2 long fwd	CTGACTCTCAGCCTGGCTAT
EPO P2 short fwd	TGGATGCTGAATGTCCTGC
EPO rev (exon 3-4 boundary)	GCTGCAGTGTTCAGCACAGC
GUSB fwd	ATCGCCATCAACAACACACT
*GUSB r*ev	TGGGATACTTGGAGGTGTCA

Cloning of *EPO* expression constructs and site-directed mutagenesis	**primer sequence 5'->3'**
EPO P1 fwd	GCAAAAGCTT**TTCCCGGGATGAGGGCC**
EPO rev	GCAAGAATTC**AGGGAGGTGGTGGATATG**
EPO exon 2 rev	GCAAGGTACC**TCTCCAGGACTCGGCT**
h*EPO* P1 ΔUTR fwd	GCAAAAGCTT**GGCGCGGAGATGGGGGTG**
h*EPO* P2 ΔUTR fwd	GCAAAAGCTT**TCTCAGCCTGGCTATCTGTT**
hEPO ΔG mutagenesis fwd	GAATGTCCTGCCTGGCTGTGCTTCTCCTG
hEPO ΔG mutagenesis rev	CAGGAGAAGCACAGCCAGGCAGGACATTC
hEPO ΔC mutagenesis fwd	GGGTGCACGAATGTCTGCCTGGCTGTGG
hEPO ΔC mutagenesis rev	CCACAGCCAGGCAGACATTCGTGCACCC

CRISPR off-target analysis	**primer sequence 5'->3'**
CRISPR OT1	TGGATGGGCGTTATTGTTGT
	GGCTGGGAGAAGGGAGGA
CRISPR OT2	GATTAGACAGGAAGTGGGG
	TGCAGTGAGCTACGATGG
CRISPR OT3	GAAGGTGAGAAGAGAAGG
	AAATAGAAGACAGGGAGG
CRISPR OT4	CCTCGTTCCTGTTAACTCT
	CACACTCCCTTCAAATCC
CRISPR OT5	CCTCTTCTCACTTTTTATCTTC
	CTCGAGAGTCCTATATTTTG
CRISPR OT6	TTTTTCTTTATTCCCAACTCC
	GCTTCCCACTATTTTTTTTCA
CRISPR OT7	TTTGTTTCTTCACCTGGC
	TCTTCATCCTTTCCCCTC
CRISPR OT8	CCTTCCCTCCAACATCCT
	TCTGTCATCCCACCACTT
CRISPR OT9	TGTGTGTATGTGTGTGTG
	ACAATTGGGGGAAAAAAAAC
CRISPR OT10	CCAGCCGTTAATATTTGT
	TCTCGTTTTTCTTTGTCTT

Supplementary Table 5. CRISPR off-target detection.

Tool	ID	Sequence	Type	Mismatch	Location	Strand	ZHANG score	COSMID score	GENE
		TCCCTGCTGTCGCTCCCTCT**GGG**	Target	0	Chr7: 100319209	+			NM_000799
			TOP 5 PREDICTED OFF-TARGETS MIT/ZHANG ALGORHITM						
ZHANG	OT1	TCCCCGCCCTCGCTCCCTCTTGG	No indel	3	Chr6: 167529802	+	1,48		intronic CCR6 (NM_004367)
ZHANG	OT2	CCCCTGCCGTCCCTCCCTCTCAG	No indel	3	Chr19: 33639182	+	1,42		intronic WDR88 (NM_173749)
ZHANG	OT3	ACCCTGCCGTCTCTCCCTCTGGG	No indel	3	Chr19: 47920846	-	1,42		intronic MEIS3 (NM_001009813)
ZHANG	OT4	TTCCTGCAGTCACTCCCTCTGGG	No indel	3	Chr11: 67038150	+	1,38		intronic ADRBK1 (NM_001619)
ZHANG	OT5	CCAATGCCGTCGCTCCCTCTGAG	No indel	4	Chr4: 96727670	-	1,37		intergenic
			TOP 5 PREDICTED OFF-TARGETS COSMID ALGORHITM						
COSMID	OT6	TCCCTGCTGTCCTTCCCTCTGGG	No indel	2	Chr4:1080908	-		1,9	intronic RNF212 (NM_194439)
COSMID	OT7	TCCCTG-TGTTGCTCCCTCTAGG	Del 14	1	Chr1:237060285	+		1,44	intronic MTR (NM_000254)
COSMID	OT8	TCCCTCCT-TCGCTCCCTCTGGG	Del 12	1	Chr12:66275392	-		1,07	intronic HMGA2 (NM_003484)
COSMID	OT9	TCCCTGCTCT-GCTCCCTCTAGG	Del 10	1	Chr19:43421026	+		1,56	intronic PSG6 (NM_021016)
COSMID	OT10	TCCCAGCTGTC-CTCCCTCTAGG	Del 9	1	Chr19:9169775	+		1,5	intergenic

List of the top five off-target sites predicted for the sgRNA used for CRISPR-mediated editing of exon 2 of the *EPO* gene by the Zhang laboratory algorithm (crispr.mit.edu) and the COSMID tool [3] (crispr.bme.gatech.edu). Red nucleotides indicate sequence differences from the target sequence. PCR amplification of the target loci followed by Sanger sequencing did not identify any off-target modifications.

Supplementary references

1. Lin M, Wei LJ, Sellers WR, Lieberfarb M, Wong WH, Li C. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. Bioinformatics 2004;20:1233-40.

2. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nature Protocols 2013;8:2281-308.

3. Cradick TJ, Qiu P, Lee CM, Fine EJ, Bao G. COSMID: A Web-based Tool for Identifying and Validating CRISPR/Cas Off-target Sites. Mol Ther Nucleic Acids 2014;3:e214.

4. Goldberg MA, Gaut CC, Bunn HF. Erythropoietin mRNA levels are governed by both the rate of gene transcription and posttranscriptional events. Blood 1991;77:271-7.

5. Uhlen M, Fagerberg L, Hallstrom BM, et al. Proteomics. Tissue-based map of the human proteome. Science 2015;347:1260419.

6. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 2011;17.

7. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15-21.

8. Aken BL, Ayling S, Barrell D, et al. The Ensembl gene annotation system. Database 2016;2016:baw093-baw.

9. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-9.

10. Lizio M, Harshbarger J, Shimoji H, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol 2015;16:22.

11. Consortium F, the RP, Clst, et al. A promoter-level mammalian expression atlas. Nature 2014;507:462-70.

12. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol 2011;29:24-6.

13. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 2011;8:785-6.

14. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. Nucleic Acids Res 2004;32:W273-9.

15. Pedersen AG, Nielsen H. Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. Proc Int Conf Intell Syst Mol Biol 1997;5:226-33.

2.2 A mutation in the Kozak sequence of the *THPO* gene causes hereditary thrombocytosis

A mutation in the Kozak sequence of the *THPO* gene causes hereditary thrombocytosis. Jakub Zmajkovic, Annalisa Pianta, Ronny Nienhold, Pontus Lundberg, Florian Groelly, Nico Ghilardi, Raoul Oude Engberink, Danielle van Lammeren-Venema, and Radek C. Skoda. Manuscript in preparation.

A mutation in the Kozak sequence of the *THPO* gene causes hereditary thrombocytosis.

Jakub Zmajkovic, Annalisa Pianta, Ronny Nienhold, Pontus Lundberg, Florian Groelly, Nico Ghilardi, Raoul Oude Engberink, Danielle van Lammeren-Venema, and Radek C. Skoda *

Word count abstract: 127 words

Word count text: 1638 words

ABSTRACT

Hereditary thrombocytosis, a myeloproliferative disorder is characterized by the elevated platelet counts. Mutations in the MPL and JAK2 genes are responsible for the primary thrombocytosis associated with low serum thrombopoietin (TPO) levels. Thrombopoietin (THPO) mutations cause secondary thrombocytosis characterized by increased TPO levels. We studied a pedigree with autosomal-dominant thrombocytosis and identified a novel activating mutation in exon 3 of the *THPO* gene. This mutation, a single nucleotide substitution G>T, is located within the Kozak sequence of the upstream open reading frame 7 (uORF7), the most critical negative regulator of *THPO* translation. We performed TPO overexpression and *in vitro* translation experiments to show that the G>T mutation disrupts the negative regulation governed by uORF7 and allows for increased translation of *THPO* protein coding sequence, ultimately causing thrombocytosis.

INTRODUCTION

Hereditary thrombocytosis also called congenital thrombocythemia is a myeloproliferative syndrome transmitted by an autosomal dominant mode of inheritance with the prevalence of < 1:100,000 in the general population.[1] It is classified by the increase in the number of thrombocytes above 450x10^9/L for the period over six months.[1,2] Based on the platelet counts, thrombocytosis may be classified as mild (450-700 x10^9/L), moderate (700-900 x10^9/L) or severe (>900 x10^9/L). [2]

Familial thrombocytosis can be classified as primary when the defect is intrinsic to the megakaryocyte progenitors in the bone marrow. Mutations in the genes involved in the thrombopoietin (TPO) signaling pathway (*MPL*[3-7] and *JAK2*[8,9]) lead to the TPO-independent signaling. Primary thrombocytosis is therefore accompanied by low TPO serum levels. Secondary thrombocytosis is caused by the mutations in the thrombopoietin (*THPO*) gene itself and is characterized by high TPO serum levels. So far, three different mutations (ΔE3, c.ΔG500, and c.G516T) in the *THPO* gene have been reported as a cause of hereditary thrombocythemia. [10-19]

THPO expression is regulated on the translational level by seven upstream open reading frames (uORF1-7) in the exons 1-3 of *THPO* mRNA, that are interfering with the translation of TPO. [20,21] uORF7 extends beyond the AUG initiating the translation of *THPO* mRNA, thus being the most critical negative regulator. All previously reported mutations in the *THPO* gene associated with hereditary erythrocytosis also destroy the negative regulation governed by uORF7. [22]

Here we report a pedigree diagnosed with thrombocytosis having a novel activating mutation in the *THPO* gene. G>T transversion weakens the Kozak consensus sequence of uORF7, thereby decreasing the translation rate of this open reading frame, which conversely allows for the increased translation of the TPO protein, leading to the thrombocytosis in the studied family.

METHODS

Study family and patient samples

The study was approved by the … and by Ethikkommission Beider Basel, Switzerland. All participating family members provided written informed consent to participate in the study. Measurements of hemoglobin, leukocytes, and platelets were carried out at the … during the time period of ….

THPO gene sequence analysis

Primers flanking exon 3 of human *THPO* (fwd: 5'-TGCCAGGCAGTCTCTTCC-3', rev: 5'-CTGCAGCGTGTCTCCTTTC-3') were used for PCR amplification with Q5 High-Fidelity DNA Polymerase (NEB). PCR products were purified with Exonuclease I and Thermosensitive Alkaline Phosphatase mix (Thermo Fisher) by incubating 15min. at 37°C followed by 85°C for 15min. Reactions were submitted for Sanger sequencing to Microsynth. Sequences were analyzed with CLC Main Workbench 7.7.1 (Qiagen).

Cell lines and transfections

HEK293LTV cell line (Cell Biolabs) was maintained in high glucose DMEM with L-glutamine supplemented with 10% fetal bovine serum (Sigma) and Penicillin/Streptomycin mix (GIBCO Invitrogen). For overexpression experiments, 2,5x105 of HEK293LTV cells per well (6-well plates) were transfected with 0.5 μg of the plasmid using jetPRIME reagent (Polyplus Transfection), and supernatants were harvested 72 hours post transfection. BaF3-hMPL cells (kindly provided by Dr. Sara C. Meyer) were cultured in RPMI media with L-glutamine, and HEPES (GIBCO Invitrogen) supplemented with 10% fetal bovine serum (Sigma), penicillin-Streptomycin mix (GIBCO Invitrogen) and 5ng/ml recombinant murine IL-3 (Peprotech). Cells were grown at 37 °C in 5 % CO2 humidified atmosphere with 21% O2.

THPO expression constructs

THPO expression constructs were prepared as described previously. [11,13] All constructs were sub-cloned into the pcDNA3 expression vector (Invitrogen). For introducing the G480T mutation into the plasmid, QuickChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies) was used with the forward 5-GCCGCCTCCATGTCCCCAGGAAGGA-3' and reverse 5'-

TCCTTCCTGGGGACATGGAGGCGGC-3' primers. Plasmid sequences were verified with Sanger sequencing (Microsynth), and midiprep DNA was produced with NucleoBond Xtra Midi kit (Macherey-Nagel).

TPO levels measurement

Cell supernatants from transfected HEK293LTV cells were sterile-filtered (0.22μm), and TPO levels were measured with human Human Tpo Quantikine Immunoassay (R&D Systems), and absorbance was measured was measured on Synergy H1 microplate reader (Biotek) equipped with Gen5 v2.05 software.

Proliferation assay

Cell proliferation rates were measured by using CellTiter-Glo Luminescent Cell Viability Assay (Promega). Supernatants from transfected HEK293LTV cells were used to stimulate the growth of BaF3-hMPL cell line. Recombinant human TPO (Peprotech) was used as a positive control. BaF3-hMPL cells were washed three times with PBS, and $5x10^3$ cells per well were plated into 96-well plates and incubated in RPMI + 10% FBS with the 10-fold dilution of cell supernatants or recombinant human TPO for 72 hours. Then, 30 μl of cells were transferred into black opaque 96-well plates (Eppendorf) and incubated with equal volume of Cell-Titer-Glo Reagent according to the manufacturer's protocol. Luminescence was measured on Synergy H1 microplate reader (Biotek) equipped with Gen5 v2.05 software.

***In vitro* transcription and translation**

THPO constructs in pcDNA3 vector were linearized with XbaI, extracted with phenol-chloroform and ethanol-precipitated. 1μg of linearized DNA was used as a template for *in vitro* transcription using T7 MEGAScript kit (Thermo Fisher Scientific). Reactions were incubated for 4 hrs at 37°C, followed by DNase treatment and phenol-chloroform extraction. The identity of the different *THPO* transcripts was verified by Agilent 2100 Bioanalyzer. 1.5ug of each *THPO* mRNA was translated for 2 hrs at 25°C using a wheat germ extract lysate (Promega) in the presence of FluoroTect Green *in vitro* Translation Labeling System (Promega). Labeled proteins were separated on NuPAGE 12% Bis-Tris Protein gels (Invitrogen) and visualized on Typhoon 9400 Imaging System (GE Healthcare Life Sciences) using 488nm laser and 526SP emission filter. Quantification of the amount of translated protein by densitometry was performed with Image Studio

Lite 5.2.5 (LI-COR Biosciences).

Statistical analysis

Statistical significance of the data was tested with one-way ANOVA or two-way ANOVA and analyzed using GraphPad Prism 7 software, P values are provided.

RESULTS

We studied the Dutch pedigree diagnosed with mild thrombocytosis (Figure 1A). Affected family members displayed elevated thrombocyte counts > 450 x10^9/L (Figure 1A and Figure 1B). Serum TPO levels stayed within the normal range of 4-31 ??/?? (Figure 1B). Sequence analysis of the *THPO* gene shown in Figure 1C revealed a heterozygous G>T transversion in the exon 3 (chr3: 184,094,114 G>T). This mutation, c.G480T was present in all affected family members and also in one family member without thrombocytosis (EN05). G480T transversion appears at the position +4 of uAUG7, which is an essential part of a Kozak consensus sequence [23] (Figure 1C). In line with this observation, translational start site score has decreased from 0.693 to 0.646 upon the introduction of the mutation, as predicted by NetStart 1.0 prediction server. [24]

We, therefore, hypothesized that G480T mutation decreases the translation initiation from uAUG7 and allows for the increased translation of *THPO* protein coding sequence. To test this, we cloned cDNAs coding for wildtype *THPO*, *THPO* with G480T mutation or other reported mutations (G516T, ΔG500, and ΔE3) into pcDNA3 vector (Figure 2A). These constructs represent the most abundant form of *THPO* mRNA that constitutes ~ 90% of *THPO* mRNAs, and uORF5 is the first ORF translated.[20] To assess the contribution of uORF7 only, we have also prepared and tested short, artificial constructs that begin with uORF7 (Supplementary Figure 1A).

Next, we transfected HEK293 cells with these constructs and measured TPO protein in cell culture supernatants by ELISA. Cells transfected with cDNAs carrying G480T mutation produced more TPO than cells transfected with wildtype construct (Figure 2B and Supplementary Figure 1B). Also, cDNAs bearing G516T, ΔG500, and ΔE3 mutations produced more TPO than the wildtype cDNA. In comparison to other mutations, G480T showed the mildest increase in TPO levels. Supernatants from cells with wildtype cDNA or cDNA carrying G480T, G516T, ΔG500, and ΔE3 mutations also stimulated the growth of TPO-dependent BaF3-MPL cells in a concentration-dependent manner, demonstrating that *THPO* G480T produces biologically active TPO protein (Figure 2C and Supplementary Figure 1C).

To directly prove that G480T increases the translational efficiency of *THPO* mRNA, we performed *in vitro* transcription-translation experiment. *In vitro* translation of the *THPO* G480T mRNA produced more TPO protein than the translation of *THPO* wildtype mRNA (Figure 2B and Supplementary Figure 1B). Also, *in vitro* translation of *THPO* G516T,

ΔG500 and ΔE3 mRNAs produced more TPO protein than *THPO* wildtype mRNA. In line with the previous reports, *THPO* mRNA with ΔE3 and ΔG500 mutations produced larger TPO molecules of ~ 41.1 kDa and ~ 40.3 kDa respectively as a result of initiation from upstream AUGs due to the in-frame fusion with uORF5 in case of ΔE3 and uORF7 in case of ΔG500 (Figure 2B and Supplementary Figure 1B).

DISCUSSION

In this study, we described a novel activating mutation in the *THPO* gene leading to the hereditary thrombocytosis. This G>T transversion resides in the exon 3 of the *THPO* gene and is a part of the Kozak consensus sequence of uORF7, a negative regulator of *THPO* translation [25] (Figure 1C). Mutations in the Kozak sequence have been previously reported to be involved in the disease pathogenesis of β-thalassemia [26], congenital heart defects [27], androgen insensitivity syndrome [28], ataxia [29], and breast cancer [30].

We demonstrated that G480T mutation destroys the negative regulation of *THPO* translation employed by uORF7. G>T transversion occurs at the position +4 of the Kozak consensus sequence that is essential for efficient translational initiation. [23] As a result, translation initiation from uORF7 is decreased, which allows ribosomes to initiate the translation from AUG8, a start codon of TPO peptide.

We showed that supernatant of HEK293 cells transfected with *THPO* G480T construct contains more TPO protein than supernatants transfected with *THPO* wildtype cDNA (Figure 2B and Supplementary Figure 1B). Cell culture supernatant is also biologically active as evidenced by the growth of TPO-dependent cell line (Figure 2C and Supplementary Figure 1C). Furthermore, G>T mutation translates into the milder phenotype than previously reported mutations (Figure 1). This observation can be explained by the fact that the guanine at +4 position is only one of the required residues of the Kozak consensus sequence. [23] In conclusion, a weak effect of G480T mutation we demonstrated in our molecular and cellular assays results in the mild thrombocytosis in the affected members of the pedigree.

FIGURE LEGENDS

Figure 1. Pedigree with hereditary thrombocytosis and identification of the *THPO* mutation. A) Schematic drawing of the pedigree. Unique patient numbers are placed under the symbols, numbers below represent the highest recorded thrombocyte values ($x10^9$/L), and the *THPO* gene sequence status is indicated as wildtype (WT), or mutated (G>T). B) Thrombocyte (PLT) and serum thrombopoietin (TPO) levels are plotted for the individuals from the pedigree shown in 1A. The grey shaded area indicates the normal range of PLT levels in the blood and TPO levels in serum. C) Chromatograms showing the DNA sequences of the mutated locus. Nucleotide sequence on the top depicts Kozak consensus sequence, the most critical residues at positions -3 and +4 marked in bold, initiating AUG triplet is framed with the pink box. The heterozygous G>T transversion in affected family members is indicated with the red box. Numbers next to the chromatograms are scores calculated by NetStart 1.0 prediction server to estimate the robustness of different AUGs to initiate the translation. The score greater than 0.5 represents a probable translation start.

Figure 2. Functional characterization of the *THPO* G>T mutation. A) Schematic drawing showing the translational reading frames of *THPO* wildtype transcript (A) and the consequence of the different mutations on the initiation of TPO translation (B-E). Open reading frames (ORFs) are depicted with numbers, red-colored are upstream ORFs and blue-colored is ORF coding for TPO protein. B) Upper part: TPO protein concentration measured in culture supernatants from HEK293 cells transfected with cDNA constructs depicted in A. TPO protein was measured by TPO ELISA and TPO concentrations are given as mean ± SD (n=3), P values were determined by one-way ANOVA. Middle part: *In vitro* transcription-translation of EPO mRNAs. Equal amounts of *in vitro* transcribed EPO mRNA variants (top) were used for *in vitro* translation in wheat germ lysates in the presence of fluorescently labeled lysine. TPO proteins separated by SDS-gel electrophoresis and visualized by fluorescent imager are shown in the bottom. The predicted molecular weight of wildtype TPO protein including signal peptide is 37.8 kDa. A higher molecular weight bands present in the lanes for ΔG500 (~ 40.3 kDa) and ΔE3 (~ 41.1 kDa) constructs represent TPO protein initiating from the upstream AUG7 that is in frame with the TPO reading frame and TPO protein initiating from the upstream AUG5, when exon 3 is skipped, respectively and are marked with an asterisk. Lower part: Densitometry of TPO protein from *in vitro* translation experiments, AU = arbitrary units.

C) The biological activity of TPO produced by transfected HEK293 cells with cDNA constructs depicted in A was measured in culture supernatants by a proliferation assay with BaF3-MPL cell line. Serial dilutions of 100ng/ml of recombinant human TPO were used for reference, n=3.

ACKNOWLEDGMENTS

We thank members of the laboratory for helpful discussions and critical reading of our manuscript. We are thankful to Sara C. Meyer for providing us with BaF3-hMPL cell line. We also thank Hui Hao-Shen, Gabi Mild and Hélène Mereau for valuable technical support.

Statement of all sources of support for our study (financial or in-kind)

This work was supported by grants from the Swiss Cancer League (KLS-2950-02-2012 and KFS-3655-02-2015), and the Swiss National Science Foundation (31003A-147016/1 and 31003A_166613) to RCS, and University of Basel "get on track" scholarship to JZ.

REFERENCES

1. Hussein K, Percy M, McMullin MF, et al. Clinical utility gene card for: hereditary thrombocythemia. Eur J Hum Genet 2014;22.

2. Teofili L, Larocca LM. Advances in understanding the pathogenesis of familial thrombocythaemia. Br J Haematol 2011;152:701-12.

3. Ding J, Komatsu H, Wakita A, et al. Familial essential thrombocythemia associated with a dominant-positive activating mutation of the c-MPL gene, which encodes for the receptor for thrombopoietin. Blood 2004;103:4198-200.

4. Moliterno AR, Williams DM, Gutierrez-Alamillo LI, Salvatori R, Ingersoll RG, Spivak JL. Mpl Baltimore: a thrombopoietin receptor polymorphism associated with thrombocytosis. Proc Natl Acad Sci U S A 2004;101:11444-7.

5. El-Harith el HA, Roesl C, Ballmaier M, et al. Familial thrombocytosis caused by the novel germ-line mutation p.Pro106Leu in the MPL gene. Br J Haematol 2009;144:185-94.

6. Liu K, Martini M, Rocca B, et al. Evidence for a founder effect of the MPL-S505N mutation in eight Italian pedigrees with hereditary thrombocythemia. Haematologica 2009;94:1368-74.

7. Teofili L, Giona F, Torti L, et al. Hereditary thrombocytosis caused by MPLSer505Asn is associated with a high thrombotic risk, splenomegaly and progression to bone marrow fibrosis. Haematologica 2010;95:65-70.

8. Bellanne-Chantelot C, Chaumarel I, Labopin M, et al. Genetic and clinical implications of the Val617Phe JAK2 mutation in 72 families with myeloproliferative disorders. Blood 2006;108:346-52.

9. Mead AJ, Rugless MJ, Jacobsen SE, Schuh A. Germline JAK2 mutation in a family with hereditary thrombocytosis. N Engl J Med 2012;366:967-9.

10. Schlemper RJ, van der Maas AP, Eikenboom JC. Familial essential thrombocythemia: clinical characteristics of 11 cases in one family. Ann Hematol 1994;68:153-8.

11. Wiestner A, Schlemper RJ, van der Maas AP, Skoda RC. An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia. Nature genetics 1998;18:49-52.

12. Kikuchi M, Tayama T, Hayakawa H, Takahashi I, Hoshino H, Ohsaka A. Familial thrombocytosis. Br J Haematol 1995;89:900-2.

13. Ghilardi N, Wiestner A, Kikuchi M, Ohsaka A, Skoda RC. Hereditary thrombocythaemia in a Japanese family is caused by a novel point mutation in the thrombopoietin gene. British journal of haematology 1999;107:310-6.

14. Graziano C, Carone S, Panza E, et al. Association of hereditary thrombocythemia and distal limb defects with a thrombopoietin gene mutation. Blood 2009;114:1655-7.

15. Kondo T, Okabe M, Sanada M, et al. Familial essential thrombocythemia associated with one-base deletion in the 5 '-untranslated region of the thrombopoietin gene. Blood 1998;92:1091-6.

16. Ghilardi N, Skoda RC. A single-base deletion in the thrombopoietin (TPO) gene causes familial essential thrombocythemia through a mechanism of more efficient translation of TPO mRNA. Blood 1999;94:1480-2.

17. Liu K, Kralovics R, Rudzki Z, et al. A de novo splice donor mutation in the thrombopoietin gene causes hereditary thrombocythemia in a Polish family. Haematologica 2008;93:706-14.

18. Stockklausner C, Echner N, Klotter AC, Hegenbart U, Dreger P, Kulozik AE. Hereditary thrombocythemia caused by a thrombopoietin (THPO) gain-of-function mutation associated with multiple myeloma and congenital limb defects. Ann Hematol 2012;91:1129-33.

19. Zhang B, Ng D, Jones C, et al. A novel splice donor mutation in the thrombopoietin gene leads to exon 2 skipping in a Filipino family with hereditary thrombocythemia. Blood 2011;118:6988-90.

20. Ghilardi N, Wiestner A, Skoda RC. Thrombopoietin production is inhibited by a translational mechanism. Blood 1998;92:4023-30.

21. Stoffel R, Wiestner A, Skoda RC. Thrombopoietin in thrombocytopenic mice: Evidence against regulation at the mRNA level and for a direct regulatory role of platelets. Blood 1995;86:3626-.

22. Cazzola M, Skoda RC. Translational pathophysiology: a novel molecular mechanism of human disease. Blood 2000;95:3280-8.

23. Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 1986;44:283-92.

24. Pedersen AG, Nielsen H. Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. Proc Int Conf Intell Syst Mol Biol 1997;5:226-33.

25. Kozak M. Emerging links between initiation of translation and human diseases. Mamm Genome 2002;13:401-10.

26. De Angioletti M, Lacerra G, Sabato V, Carestia C. β+45 G → C: A novel silent β-thalassaemia mutation, the first in the Kozak sequence. British journal of haematology 2004;124:224-31.

27. Mohan RA, van Engelen K, Stefanovic S, et al. A mutation in the Kozak sequence of GATA4hampers translation in a family with atrial septal defects. American Journal of Medical Genetics Part A 2014;164:2732-8.

28. Choong CS, Quigley CA, French FS, Wilson EM. A novel missense mutation in the amino-terminal domain of the human androgen receptor gene in a family with partial androgen insensitivity syndrome causes reduced efficiency of protein translation. J Clin Invest 1996;98:1423-31.

29. Usuki F, Maruyama K. Ataxia caused by mutations in the alpha-tocopherol transfer protein gene. J Neurol Neurosurg Psychiatry 2000;69:254-6.

30. Signori E, Bagni C, Papa S, et al. A somatic mutation in the 5'UTR of BRCA1 gene in sporadic breast cancer causes down-modulation of translation efficiency. Oncogene 2001;20:4596-600.

Figure 1

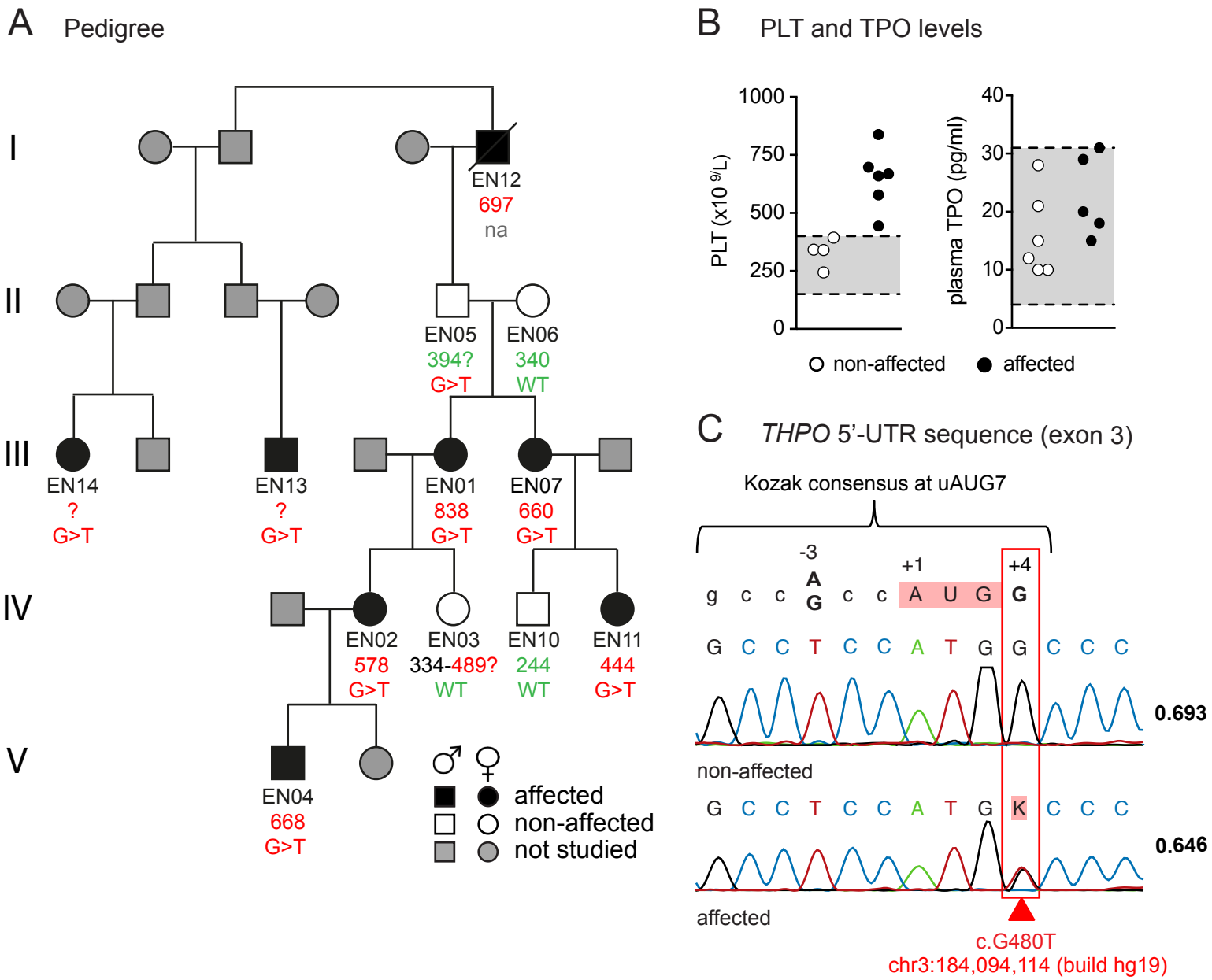
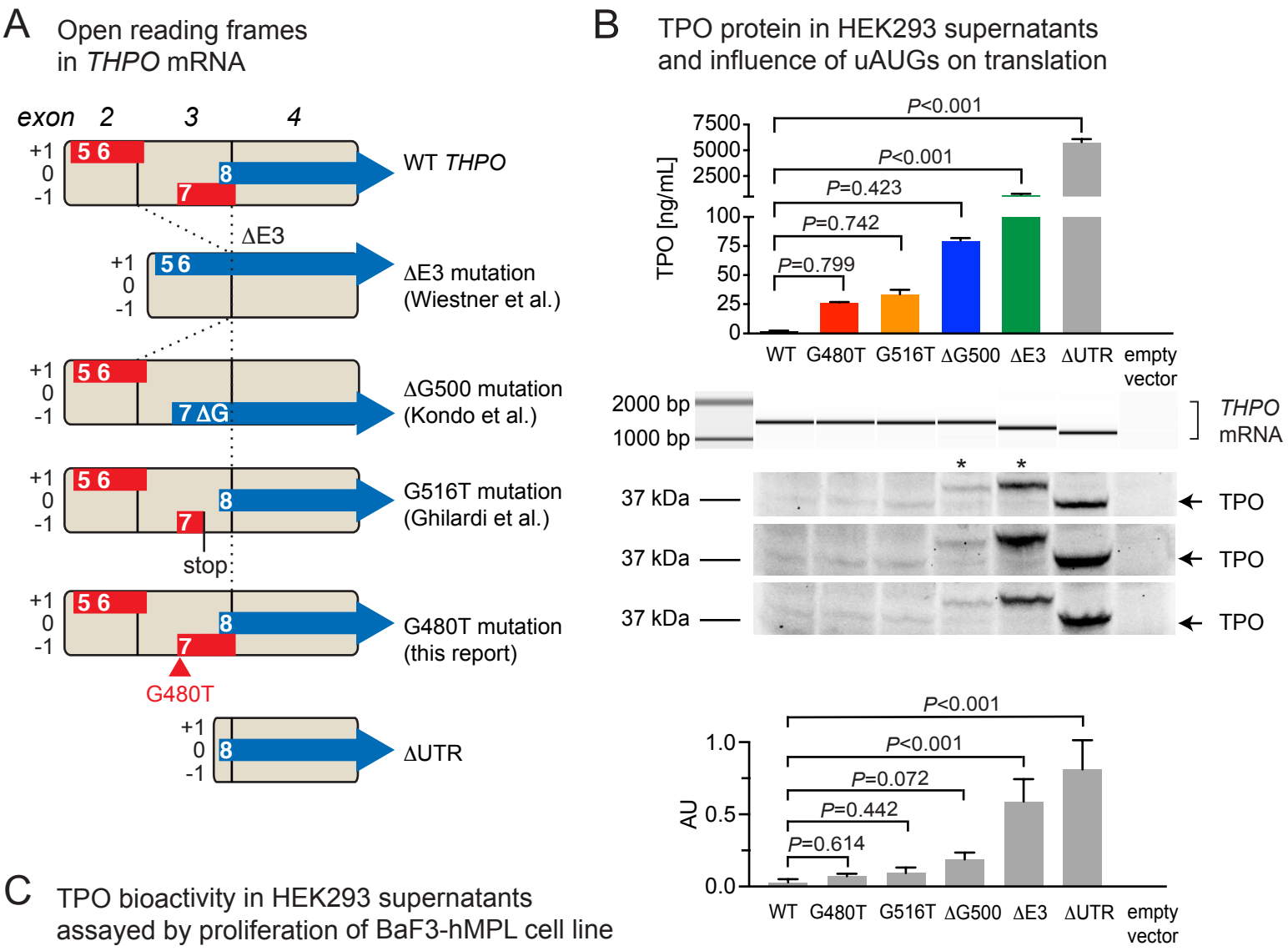
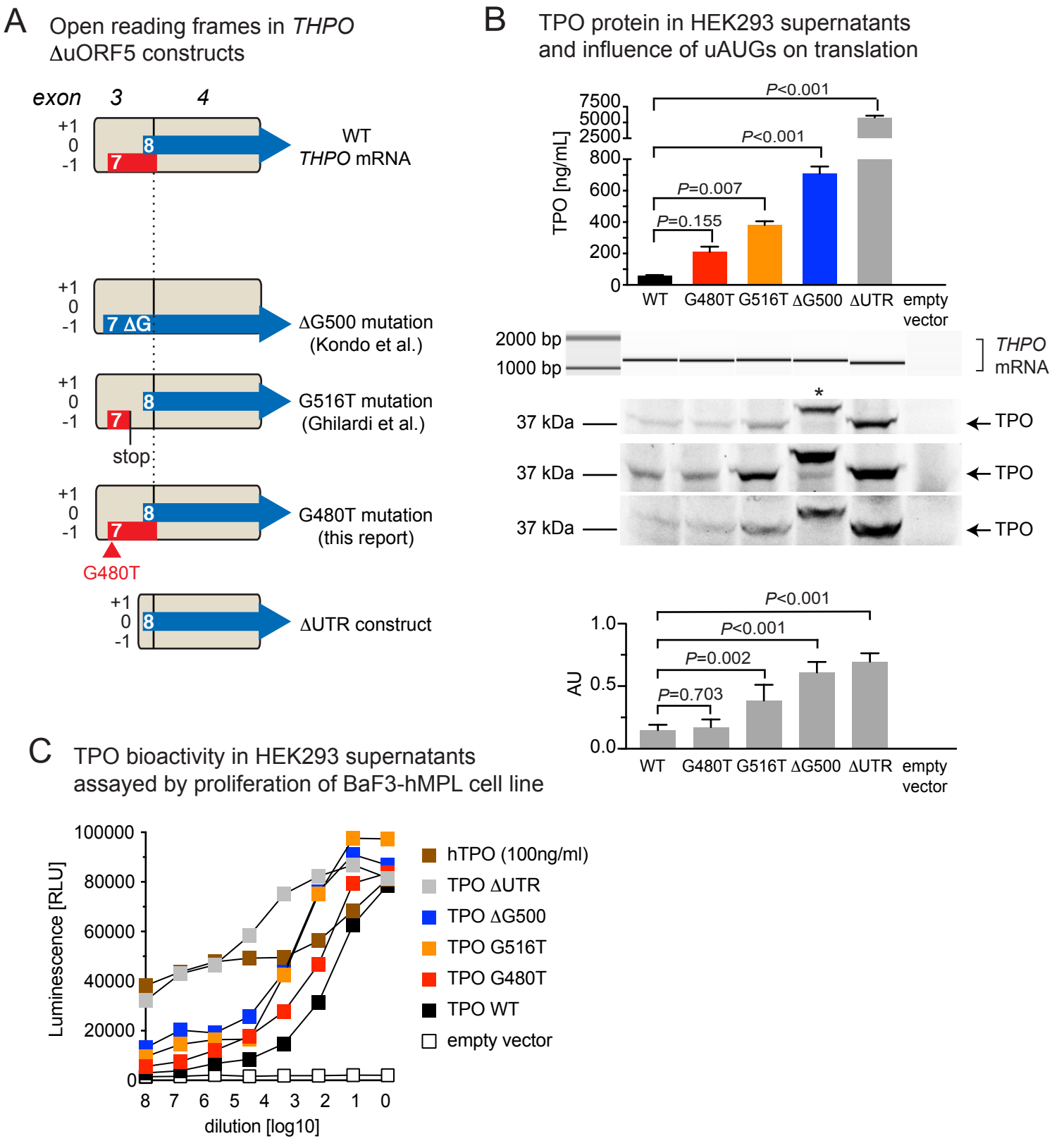

A
Pedigree
I
II
III
IV
V
EN12
697
na
EN05
394?
G>T
EN06
340
WT
EN14
?
G>T
EN13
?
G>T
EN01
838
G>T
EN07
660
G>T
EN02
578
G>T
EN03
334-489?
WT
EN10
244
WT
EN11
444
G>T
EN04
668
G>T
affected
non-affected
not studied
B
PLT and TPO levels
PLT (x10 9/L)
1000
750
500
250
0
plasma TPO (pg/ml)
40
30
20
10
0
non-affected
affected
C
THPO 5'-UTR sequence (exon 3)
Kozak consensus at uAUG7
-3
+1
+4
g c c A G c c A U G G
G C C T C C A T G G C C C
0.693
non-affected
G C C T C C A T G K C C C
0.646
affected
c.G480T
chr3:184,094,114 (build hg19)

Figure 2


A
Open reading frames in THPO mRNA
exon 2 3 4
+1 0 -1
5 6
7
8
WT THPO
ΔE3
ΔE3 mutation (Wiestner et al.)
7 ΔG
ΔG500 mutation (Kondo et al.)
G516T mutation (Ghilardi et al.)
stop
G480T mutation (this report)
G480T
ΔUTR
B
TPO protein in HEK293 supernatants and influence of uAUGs on translation
TPO [ng/mL]
P<0.001
P<0.001
P=0.423
P=0.742
P=0.799
WT G480T G516T ΔG500 ΔE3 ΔUTR empty vector
2000 bp
1000 bp
THPO mRNA
37 kDa
TPO
AU
P<0.001
P<0.001
P=0.072
P=0.442
P=0.614
C
TPO bioactivity in HEK293 supernatants assayed by proliferation of BaF3-hMPL cell line
Luminescence [RLU]
dilution [log10]
hTPO (100ng/ml)
TPO ΔUTR
TPO ΔE3
TPO ΔG500
TPO G516T
TPO G480T
TPO WT
empty vector

Supplementary Figure 1



Supplementary Figure 1. Functional characterization of the *THPO* G>T mutation with ΔuORF5 constructs. A) Schematic drawing of *THPO* constructs lacking uORF5 analogous to the ones in Figure 2. B-C) TPO ELISA and in *vitro* transcription-translation (B), and proliferation assay (C) were performed with ΔuORF5 constructs analogous to the Figure 2. Note that the results of the ΔUTR construct, the empty vector and hTPO control shown for reference are taken from Figure 2C.

2.3 Murine models to study MPN

2.3.1 Lineage bias of HSPC population

One of the crucial questions in the MPN field is how a single mutation in JAK2 (*JAK2 V617F*) can cause clinically distinct diseases (PV, ET, and PMF). Recently, multiple reports described the heterogeneity of HSC compartment (Perié et al., 2015; Sanjuan-Pla et al., 2013; Yamamoto et al., 2013) with the possible bias towards a particular lineage, demonstrated for the megakaryocyte-biased HSCs (Grover et al., 2016; Haas et al., 2015; Sanjuan-Pla et al., 2013). These HSCs are thought to predominantly give rise to a special type of progenitor, and mutational hit in such a stem cell could explain different phenotype outcomes. To investigate this in more detail, we plan to FACS sort "biased" stem cells with a combination of different cell surface markers, and to subsequently transplant them into lethally irradiated BL6 recipients.

Standard transplantation procedure in the lab involves transgene induction (SclCre; *JAK2 V617F*) in the donor mice followed by FACS sorting of pre-induced (mutated) cells and subsequent transplantation. The caveat of this setup is, however, that the expression pattern of some of the markers used for cell sorting may be changed upon *JAK2 V617F* activation (Lundberg et al., 2014). Moreover, unpublished results from our lab suggest, that whole bone marrow cell suspension is more effective in disease initiation than the population of sorted HSCs. This could have three possible explanations:

1. cells in the bone marrow of non-hematopoietic origin are playing an active part in the disease initiation and progression.
2. mechanical stress during FACS sorting damages LT-HSC and impair their proliferation capacity.
3. FACS sorting of pre-induced HSC does not enrich for the correct cell population.

To study the third possibility, we attempted to set-up a system allowing us to sort non-induced HSCs, and activate them afterward, either *in vitro* or *in vivo*. For the *in vitro* part, we isolated bone marrow cells from *JAK2 V617F* donor mice (Tiedt et al., 2008) and FACS-sorted LT-HSC population (defined as Lin^-, $Sca\text{-}1^+$, $c\text{-}kit^+$, $CD150^+$, $CD48^-$ population). Next, LT-HSCs were incubated with different concentrations of 4-hydroxytamoxifen (4-OHT) to activate the transgene and plated on methyl-

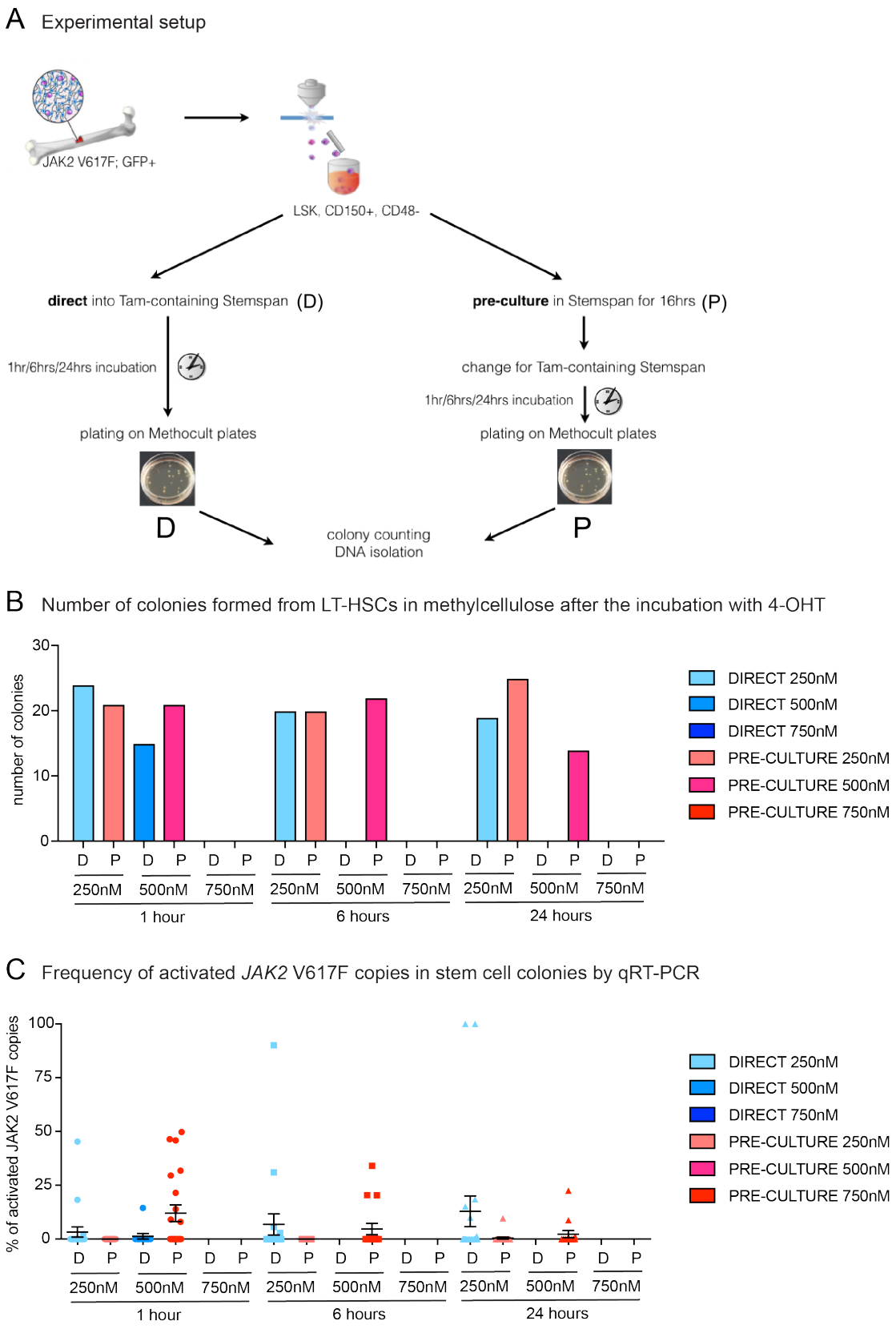


Figure 13. Activation of LT-HSC by 4-OHT *in vitro*. A) Experimental setup describing the isolation of LT-HSC population (defined as Lin^-, $Sca\text{-}1^+$, $c\text{-}kit^+$, $CD150^+$, $CD48^-$ subset) and subsequent incubation with 4-OHT (Tam), either in direct culture (D) or with the pre-culture in Stemspan medium (P) for various time points. Cells were then plated on the methylcellulose plates, let to grow colonies, picked, DNA was isolated, and the number of activated transgene copies was assessed by qRT-PCR. B) The number of colonies grown from LT-HSCs incubated for different time points with varying concentrations of 4-OHT. C) Percent of activated copies of the human JAK2 V617F transgene.

cellulose plates to grow colonies (Figure 13A). Prior to the 4-OHT exposure, cells were either preincubated for 16 hours in Stemspan (P), or directly sorted into the Stemspan with 4-OHT (D). We hypothesized that pre-culture in Stemspan might help LT-HSCs to enter the cell cycle and be more accessible for Cre-mediated recombination. Afterward, DNA was isolated from colonies and subjected to the qRT-PCR to assess the number of activated transgene copies in single colonies.

Colony counting showed, that both high concentrations of 4-OHT (750nM) and prolonged incubation times had a negative impact on the colony survival. Conversely, pre-culture of LT-HSCs prior to the 4-OHT exposure led to the increased colony numbers (Figure 13B). Transgene activation in LT-HSCs occurred at a very low frequency as revealed by qRT-PCR analysis. The best results were achieved with pre-incubation in Stemspan for 16 hours followed by 1-hour exposure to 500nM 4-OHT (Figure 13C). Low efficiency could be explained by the fact, that Cre recombinase activity is cell-cycle dependent (mainly in late S to M phase) (Jo et al., 2003), whereas LT-HSCs are mainly dormant and arrested in G0 phase (Nakamura-Ishizu et al., 2014; Wilson et al., 2008). Live cell imaging demonstrated, that the first division of LT-HSC *in vitro* takes on average 45-50 hours (Haetscher et al., 2015). In conclusion, we find *in vitro* approach not feasible due to the low efficiency and technical limitations.

For the *in vivo* part, we again FACS sorted non-induced LT-HSCs from donor mice and transplanted them into lethally irradiated BL6 recipients, that were either pre-treated with tamoxifen at various time, or their diet was supplemented with tamoxifen-containing food (Figure 14A) (Reinert et al., 2012; Wilson et al., 2014). Follow-up of transplanted mice by measuring serial blood counts and GFP chimerism revealed, that some of the mice developed MPN. Disease initiation positively correlated with increasing tamoxifen dose and time course of the administration. However, high tamoxifen dose also had a negative effect on survival (data not shown). After several trials, we established the final dose that induces HSCs *in vivo* with sufficient efficiency and does not cause any lethality: a single dose of 5mg of tamoxifen administered at the time of transplantation (Figure 14B, green group). Although tamoxifen-supplemented diet also showed high efficiency, it could potentially lead to the high intergroup variabilities. We plan to use this setup in elucidating the role of biased stem cells and bone marrow microenvironment in disease initiation and progression.

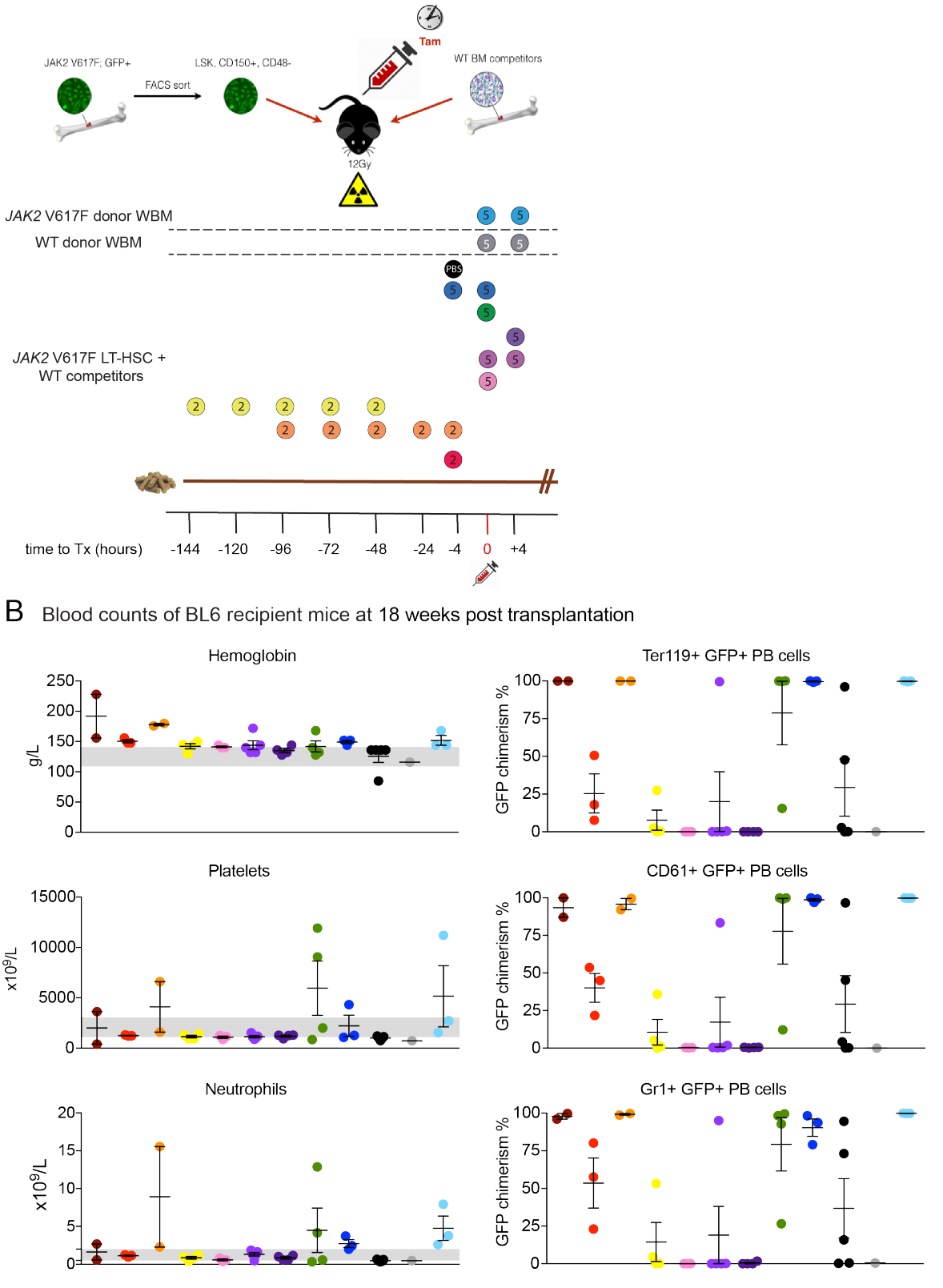


Figure 14. Activation of LT-HSC by 4-OHT *in vivo*. A) Experimental setup describing the isolation of LT-HSC population (defined as Lin$^-$, Sca-1$^+$, c-kit$^+$, CD150$^+$, CD48$^-$ subset) and subsequent transplantation into tamoxifen-pre-treated lethally irradiated recipients together with wildtype whole bone marrow competitors. Each circle represents a single tamoxifen injection with the number indicating the tamoxifen dose in milligrams. Color corresponds to the different treatment groups displayed in B. Axis illustrates the time of tamoxifen dosing with respect to the transplantation that is displayed as 0. B) Blood parameters and respective chimerism at week 18 post transplantation. The grey-shaded area represents the normal reference range. Colors for different groups correspond to different treatment regimens displayed in A.

2.3.2 Tracking MPN initiation *in situ*

Our lab previously showed that a single *JAK2 V617F* hematopoietic stem cell is capable of initiating and driving the MPN disease (Lundberg et al., 2014). However, the majority of studies in the hematopoietic stem cell field are using transplantation of cells into lethally irradiated animals. This procedure wipes out the recipient's bone marrow niche and creates cytokine release syndrome, that favors the engraftment of transplanted stem cells and possibly overestimates the disease-initiating capacity of *JAK2 V617F* stem cells.

To overcome this limitation, we used Confetti reporter system that allows for tracking of clonal stem cell origin *in situ*. Confetti mouse strain has a transgene cassette knocked-in in Rosa26 locus under the control of Cre recombinase. This cassette contains 4 fluorophores (CFP-GFP-YFP-RFP) that are randomly activated upon Cre administration (Schepers et al., 2012; Snippert et al., 2010). In a homozygous state, the cassette can produce up to 10 different color combinations (Baggiolini et al., 2015). Crossing this strain with our *JAK2 V617F* model allows tagging the disease-initiating cells fluorescently. We aimed for a limiting dose of tamoxifen to induce only a few stem cells to have each stem cell labeled with a different color. Double transgenic mice *SclCre*^{ER}; *JAK2 V617F*$^{fl/+}$; Confetti$^{fl/fl}$ were induced by single injection with various tamoxifen dosage.

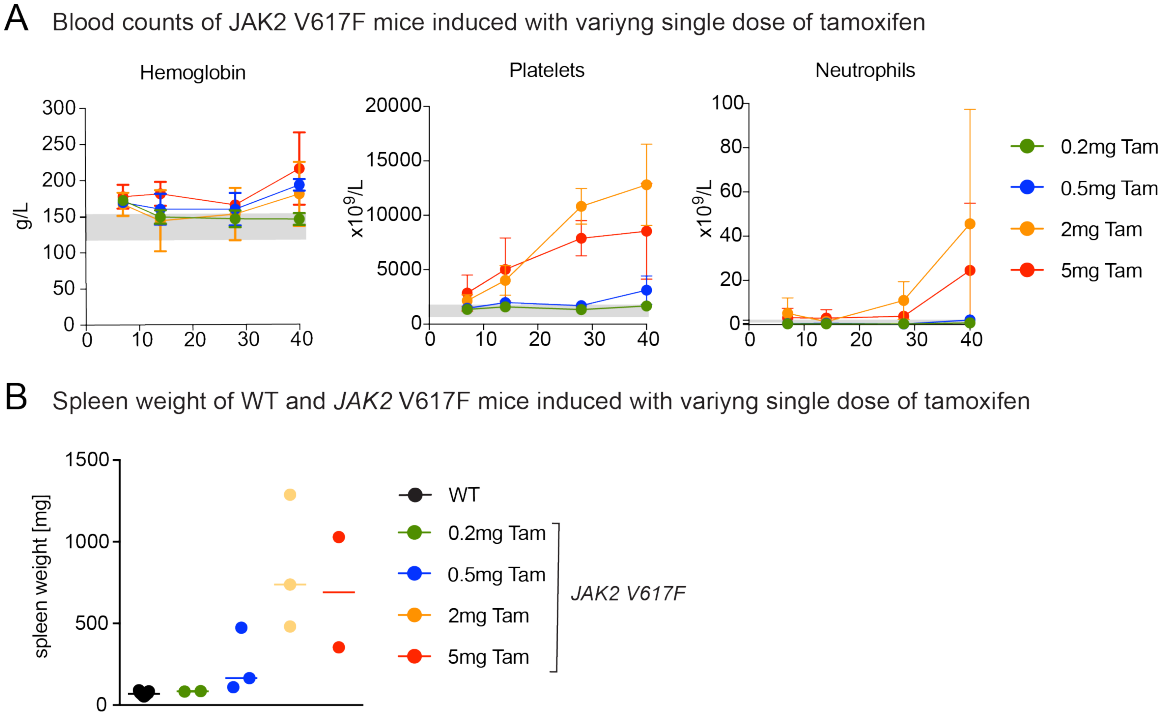


Figure 15. Blood parameters and spleen weight of *SclCre*^{ER}; *JAK2 V617F*$^{fl/+}$; Confetti$^{fl/fl}$ mice. Blood parameters in A) and spleen weight in B) of *SclCre*^{ER}; *JAK2 V617F*$^{fl/+}$; Confetti$^{fl/fl}$ mice treated with a single tamoxifen injection. The grey-shaded area represents the normal reference range.

We monitored the disease progression by measuring serial blood counts, composition and frequency of different fluorescently labeled clones by flow cytometry. We could observe a dose-response, as mice with the highest tamoxifen dose developed the most severe disease (Figure 15A and 15B) and had the highest proportion of cells fluorescently tagged (Figure 16A). Interestingly, increasing the dose from 2mg to 5mg of tamoxifen did not lead to a further increase in the number of activated cells (Figure 16A). Since we were able to detect labeled cells 40 weeks after the tamoxifen dosing, we assume that also long-term repopulating stem cell and progenitor compartment was targeted by the tamoxifen induction (Figure 16A-C).

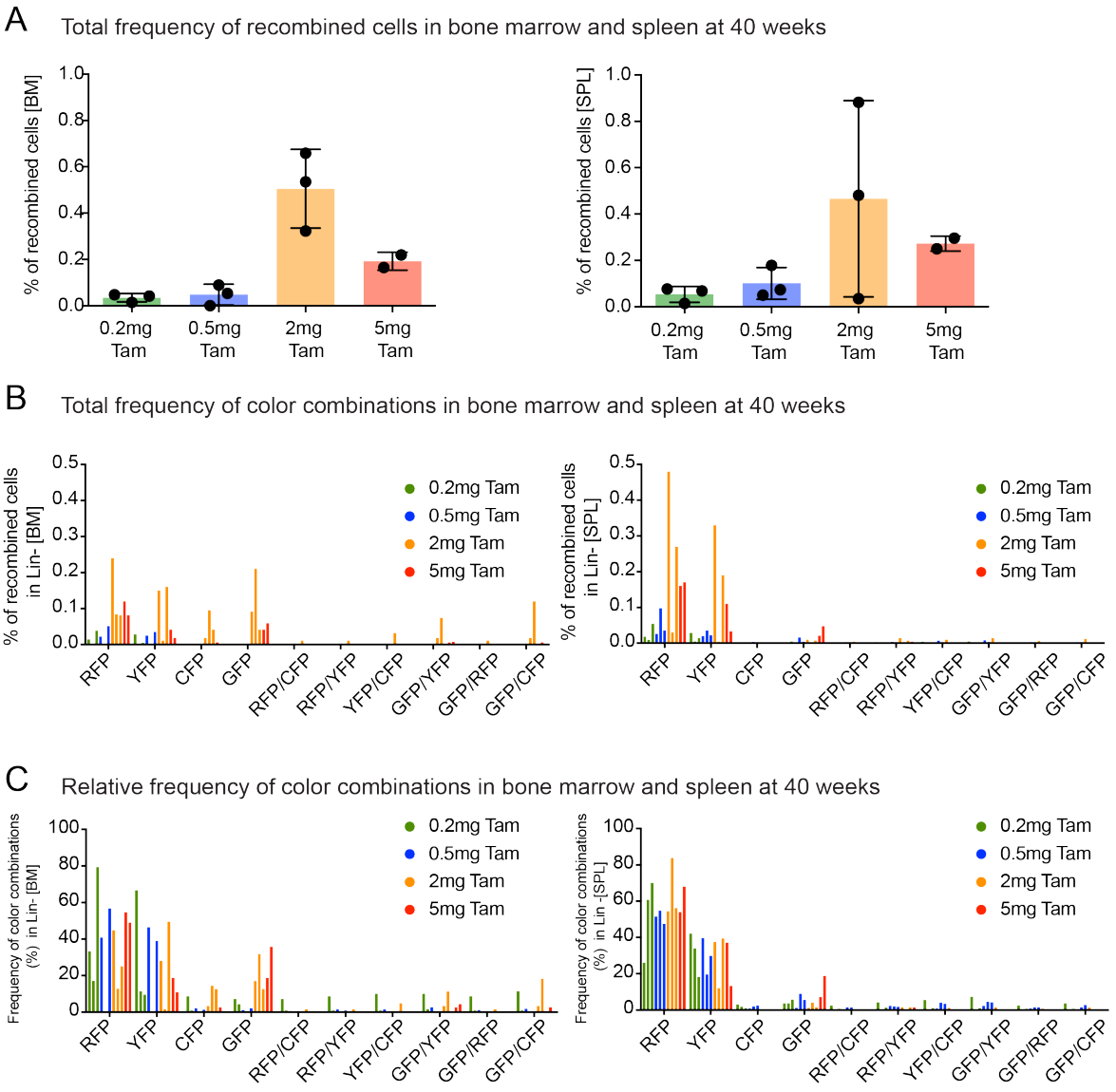


Figure 16. The frequency of different color combinations in *SclCre*^{ER}; *JAK2 V617F*$^{fl/+}$; Confetti$^{fl/fl}$ mice 40 weeks after the tamoxifen injection of Lin^{-} subset in bone marrow (BM) and spleen (SPL). The total frequency of recombined cells (A), the total frequency of color combinations (B) and the relative frequency of color combinations (C) is shown. In A, an average of the group (2-3 mice) is plotted, whereas in B and C each bar represents an individual mouse.

Surprisingly, we observed a bias in fluorophore activation, as not all colors were detected with a similar frequency. RFP and YFP represented ~ 70% of activated cells, and single colors were activated with higher efficiency than the double color combinations (Figure 16B and 16C). This phenomenon was reported before and currently, there is no explanation for it (Baggiolini et al., 2015). More importantly, we could detect human JAK2 expression (transgenic V617F allele) in the fluorescent-negative subset of Lin- population, meaning, that 2 transgenes we combined in this mouse model have a different threshold of activation (Figure 17). JAK2 V617F transgene being more sensitive to tamoxifen yields cells with activated human JAK2, but without a fluorescent label. Two abovementioned observations demonstrate that the tested model is not suitable for intended studies of MPN initiation from single HSPC cell. Thus, different approaches, such as molecular barcoding need to be explored.

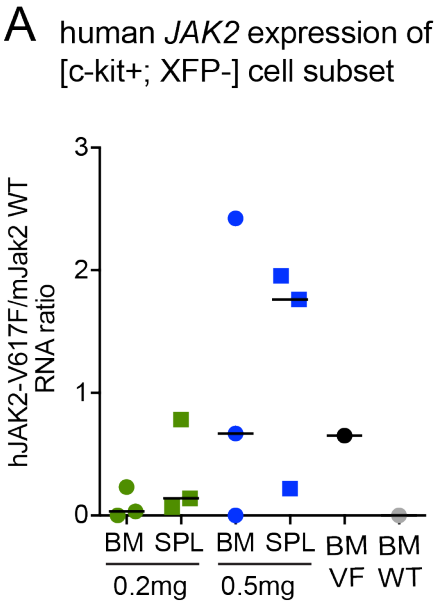


Figure 17. Human JAK2 expression in c-kit^{+}; XFP^{-} compartment. qRT-PCR analysis of human JAK2 expression in bone marrow (BM) and spleen (SPL) of mice treated with low tamoxifen dose together with the bone marrow of JAK2 V617F mouse (BM VF) and wildtype mouse (BM WT) that were used as controls.

3 Discussion

3.1 *EPO* gene mutation and erythrocytosis

Erythropoietin is to date the only identified humoral factor regulating erythropoiesis. Therefore, mutations in the gene coding for *EPO* are thought to disturb this regulation. Indeed, Arg150Gln substitution leading to loss-of-function and causing Diamond-Blackfan-like anemia was reported (Kim et al., 2017). Conversely, gain-of-function mutations in the *EPO* gene causing erythrocytosis have not been described yet.

Earlier this year we described a novel gain-of-function mutation in the *EPO* gene causing an inherited form of erythrocytosis and elucidated the molecular mechanism of the disease pathogenesis (Zmajkovic et al., 2018). We identified the candidate mutation based on the genetic linkage with a LOD score of 3.3 for the locus on chromosome 7q21, in a pedigree with 10 affected family members. This mutation, a heterozygous deletion of a G nucleotide at a position 32, introduces a frameshift in exon 2 that interrupts translation of the main *EPO* mRNA (P1 transcript). It appeared to be a loss-of-function mutation at first sight. However, it also initiates excess production of erythropoietin from alternative non-coding mRNAs (P2 transcripts) located in intron 1 of the *EPO* gene. P2 transcripts produce biologically active EPO due to the frameshift inflicted by ΔG that provides a bridge to the open reading frame coding for physiological EPO. Therefore, mutations in EPO should be considered in the search for causes of secondary erythrocytosis. Finally, our study also demonstrates the usefulness of CRISPR/Cas9 genome editing in elucidating the molecular mechanisms of rare diseases.

P2 transcripts were the predominant mRNA species in the liver, whereas P1 transcripts were mainly detected in RNA from the kidney. Thus, we suspect that the excessive amounts of EPO in the affected family members are coming from the liver. Glycosylation of EPO is critical for maintaining longer half-life in the serum (Goldwasser et al., 1974). A method to determine the overall glycosylation pattern of EPO, based on the chromatographic separation of the EPO isoforms was developed (Lönnberg et al., 2012). Intriguingly, EPO was significantly more glycosylated in the human umbilical cord plasma samples than in the samples from the peripheral blood (de Seigneux et al., 2016). Since EPO in the human umbilical

cord plasma is of the hepatic origin and EPO in the peripheral blood plasma of a renal origin, it seems that these two organs glycosylate EPO differently. Thus, it would be interesting to measure the glycosylation pattern of EPO in affected and non-affected family members, as we would expect that affected family members with the majority of EPO produced in the liver would have higher glycosylation level compared to the non-affected family members.

Erythropoietin executes a fundamental role in the organism, and this importance is also reflected in the conservation of its mRNA sequence among the species. Noteworthy, there is also conservation of certain non-coding regions of EPO between human and mouse, including the intron 1, with overall homology over 65% (Shoemaker and Mitsock, 1986). It was suggested that this region might exhibit a regulatory function (Imagawa et al., 1991; Madan et al., 1995; Shoemaker and Mitsock, 1986).

Interestingly, gene regulatory elements are mostly found within the 5'-most introns (Beaulieu et al., 2011; Bianchi et al., 2009; Coulon et al., 2010; Gaunitz et al., 2005; 2004; Kaneko et al., 2014; Ott et al., 2009; Scohy et al., 2000; Tourmente et al., 1993). Furthermore, intronic RNAs constitute the major fraction of the non-coding RNA and contain a large number of potential endogenous siRNAs (Rearick et al., 2011; St Laurent et al., 2012).

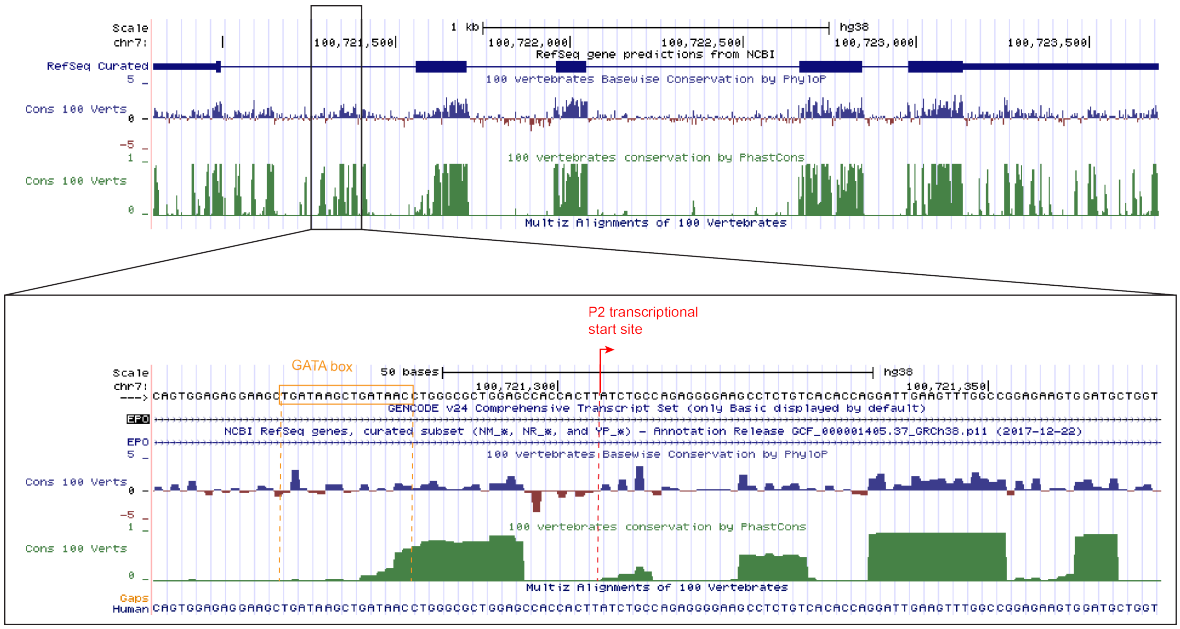


Figure 18. Intron 1 of the *EPO* gene is conserved amongst vertebrates. The upper panel shows UCSC tracks PhyloP and PhastCons displaying the conservation of *EPO* genomic sequence amongst 100 vertebrates. The lower panel is a zoom in the region of intron 1, where GATA boxes within the putative promoter in orange and P2 transcriptional start side in red reside.

The physiological function of wild-type P2 transcripts is currently unknown, but it seems likely that they do have a function because one of the conserved regions within intron 1 contains P2 transcriptional start site (Figure 18). We found that the long form of *EPO* P2 transcript is vastly overexpressed in the adult human liver compared to the kidney. The liver is considered to be the primary organ of the EPO production in the embryonic development. Around the birth, EPO production gradually moves to kidneys that become the primary site of the production of EPO in adult life. Although this switch is well documented in multiple species (Bondurant et al., 1991; Dame et al., 1998; Eckardt et al., 1992; Peschle et al., 1975; Zanjani et al., 1981), molecular mechanisms leading to this switch remain elusive. It was suggested, that the spatio-temporal expression of GATA transcription factors may play a role in this process (Weidemann and Johnson, 2009).

GATA proteins share related Cys-X_2-Cys-X_{17}-Cys-X_2-Cys (where X represents any amino acid residue) zinc-finger DNA-binding domains among mammals that bind to the consensus motif 5′-(A/T)GATA(A/G)-3′ (Simon, 1995) (Simon, 1995), and play critical roles in cell growth and differentiation (Lentjes et al., 2016). In the core promoter region of erythropoietin, a GATA factor-binding motif (GATA box) with the core sequence A**GATAAC**A resides. Intron 1 putative promoter also contains two consecutive GATA boxes T**GATAAG**CT**GATAAC**C. It was reported that human GATA-1, GATA-2, and GATA-3 are able to bind to the core *EPO* promoter sequence, inhibit the formation of pre-initiation complexes, and thereby reduce the expression of *EPO* mRNA in Hep3B cells (Aird et al., 1994; Imagawa et al., 1997; 2002). In addition, GATA-2 and GATA-3 were demonstrated to constitutively repress *EPO* expression in renal tubular cells in the kidney *in vivo* (Obara et al., 2008). Thus, it was proposed that GATA box acts as a negative regulatory element, repressing the *EPO* transcription during normoxia. Under hypoxic conditions, GATA binding markedly decreases, which allows for a significant increase in *EPO* gene expression. In contrast, GATA-4 was proposed to have a positive effect on *EPO* expression in Hep3B cells, since knockdown of GATA-4 negatively regulated *EPO* mRNA levels (Dame et al., 2004). We looked at the distribution of GATA transcripts in human kidney and liver samples in the publicly available RNASeq database of ProteinAtlas (Uhlen et al., 2017) and also analyzed GATA expression in kidney and liver specimens by qRT-PCR (Figure 19).

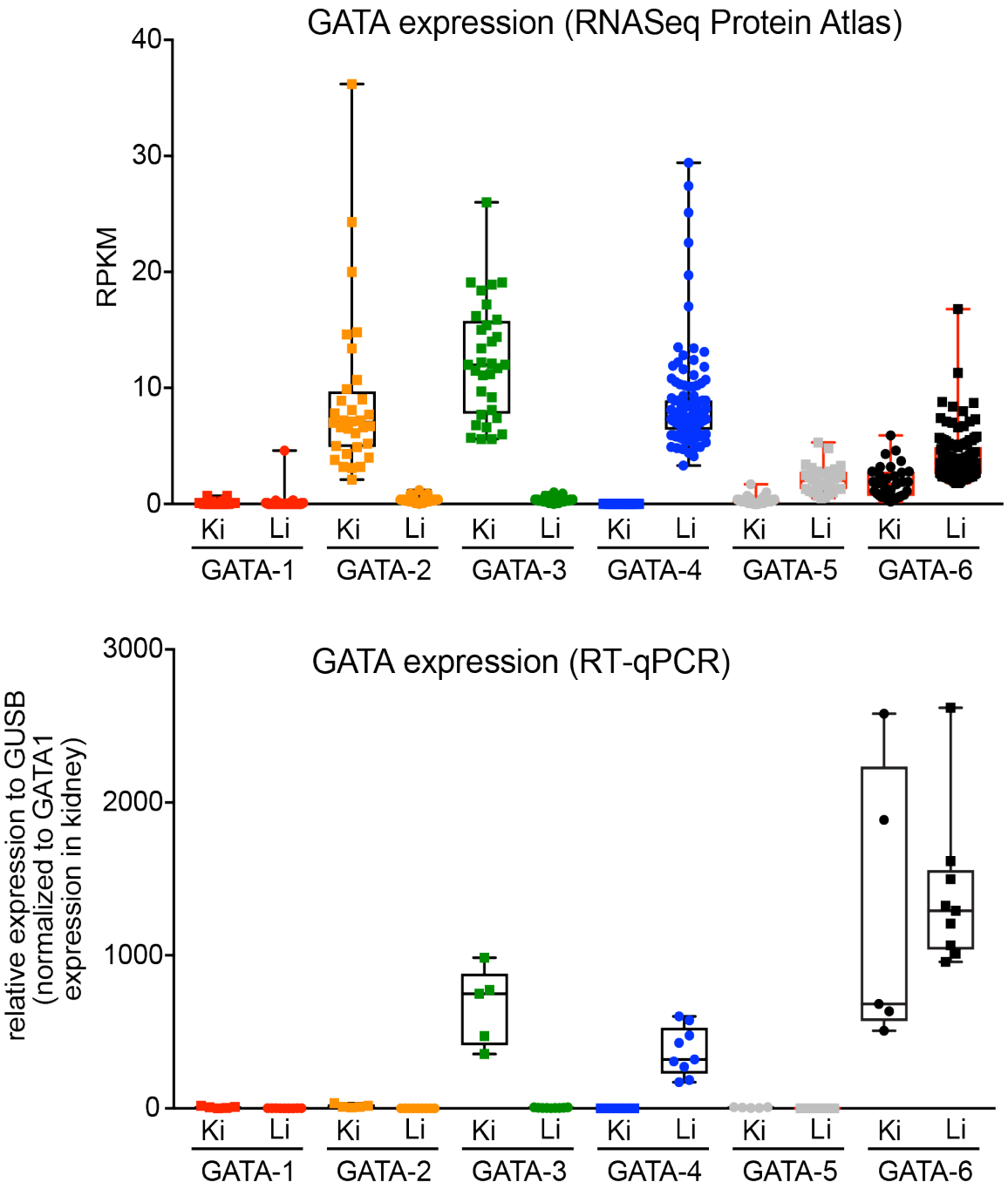


Figure 19. The upper graph shows the expression of GATA proteins in human kidney and liver as determined by RNAseq (ProteinAtlas). The lower graph displays the expression of GATA proteins in human kidney and liver as determined by qRT-PCR.

In both cases, we found that GATA-3 and GATA-4 showed tissue-specific expression pattern. GATA-3 was expressed in the kidney and not in the liver, whereas GATA-4 was expressed in the liver, but not in the kidney. In line with this observations, GATA-3 knockout in the mouse is embryonically lethal at day E12.5, due to the aberrations in fetal liver hematopoiesis (Pandolfi et al., 1995) and renal morphogenesis defects (Grote et al., 2006). Furthermore, individuals with GATA-3 haploinsufficiency suffer from renal anomalies as a part of HDR Syndrome

(Hypoparathyroidism, Sensorineural Deafness and Renal Disease) (Van Esch et al., 2000). Disruption of GATA-4 also leads to the embryonic lethality at day E9.5 due to the heart morphogenesis defects and defects in the development of embryonic liver (Kuo et al., 1997; Watt et al., 2004).

Based on the evidence above, we hypothesize that *EPO* P2 transcripts are silencing the expression of P1-coded erythropoietin in hepatocytes after birth. Liver-restricted expression of P2 transcripts may be positively regulated by GATA-4 or negatively regulated by GATA-3 transcription factors. To get a deeper understanding of this regulation, we aimed to generate a Hep3B cell line, knocked-out for *EPO* P2 transcripts, by employing the CRISPR/Cas9 genome editing approach. We pursued to knock-out two GATA-boxes of the P2 putative promoter in the intron 1. The schematics of targeting of the locus by one of gRNAs is shown in Figure 20.

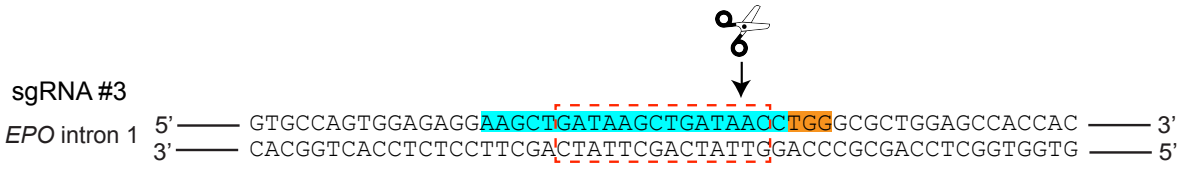


Figure 20. An example of gRNA targeting GATA boxes located within the intron 1 of *EPO*. Single-guide RNA targeting sequence is depicted in light blue and protospacer adjacent motif (PAM) in orange. The arrow shows the cut site. GATA boxes are framed in red.

We were able to generate multiple single-cell clones with different indels in the GATA-box region (Figure 21). We first plan to confirm that the *EPO* P2 putative promoter knock-out clones do not express P2 transcripts. Next, to test our hypothesis, that *EPO* P2 transcripts might be responsible for silencing the expression of *EPO* in hepatocytes, we will measure EPO production in cell supernatants by ELISA. If our hypothesis is correct, we expect that the expression of EPO from the P1 promoter will increase in these cells.

It could also well be, that the P2 transcripts regulate the function of other genes in the form of a long non-coding RNA, or a short peptide produced by the translation of P2 transcript (Anderson et al., 2015; Kondo et al., 2010; Magny et al., 2013; Nelson et al., 2016; Pauli et al., 2014). If we do not see an effect on the erythropoietin production, we aim to identify other potential targets of P2 transcripts by RNAseq. To elucidate the potential contribution of GATA proteins in the spatially-restricted expression of P2 transcripts, we will perform knock-down of GATA-4 by siRNA and overexpression of GATA-3 in Hep3B cell line. We also plan to perform

CHIP experiments to identify, whether GATA proteins bind to the putative promoter in intron 1 of the *EPO* gene.

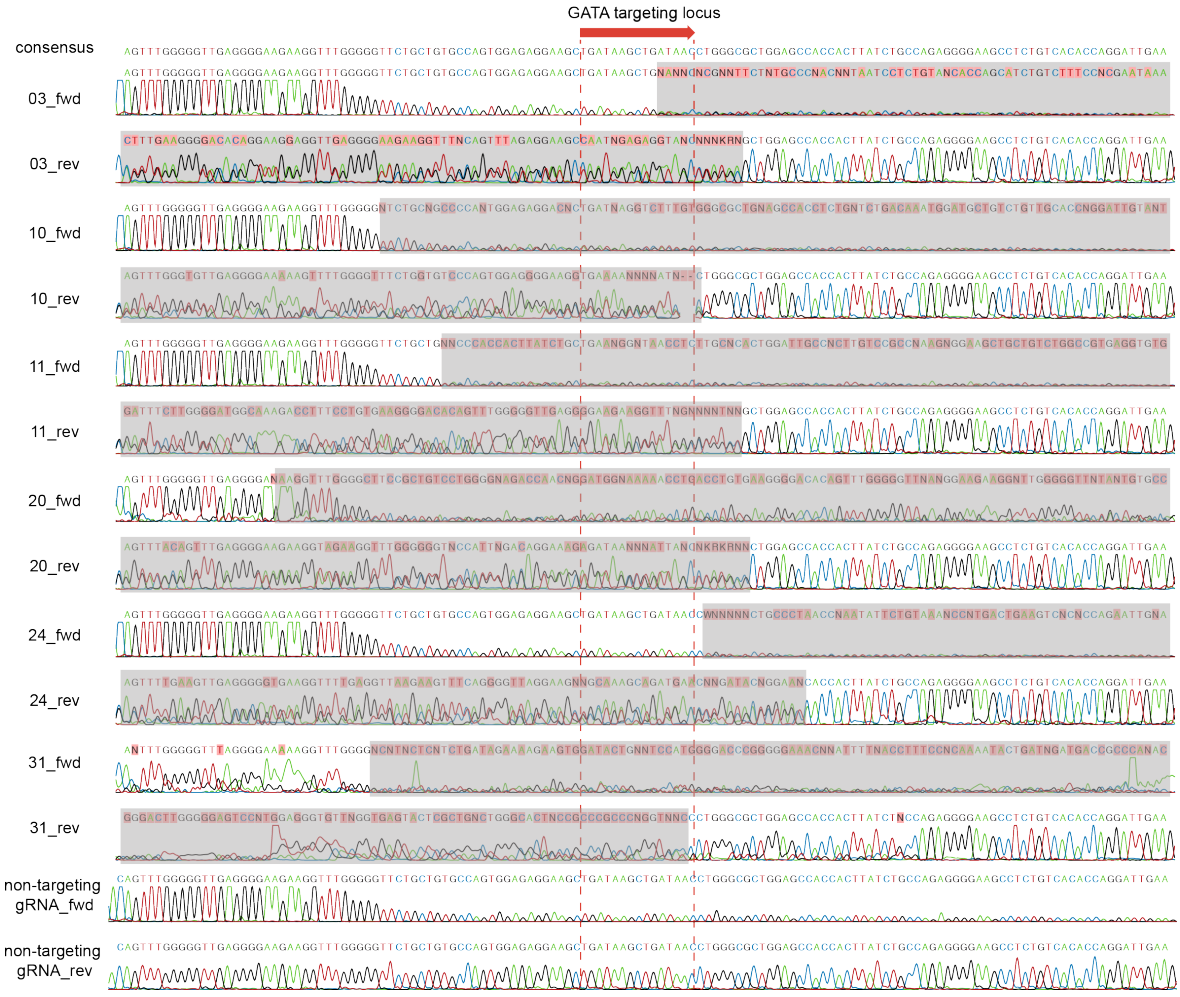


Figure 21. Sequencing chromatograms of Hep3B single cell clones edited with CRISPR/Cas9. Grey-shaded areas mark the out-of-frame sequence due to the indels generated by CRISPR/Cas9. A red arrow with red dotted lines marks the GATA box region. Both forward and reverse reads are displayed for each clone. Non-targeting gRNA is a control gRNA with the random sequence that is not compatible with any region in the human genome.

3.2 *THPO* gene and thrombocytosis

Thrombopoietin is the primary cytokine regulating the proliferation of megakaryocyte progenitors expressed by liver hepatocytes. Until recently, TPO production was sought to be constitutive (Stoffel et al., 1995), with TPO plasma levels regulated by the uptake of the TPO by megakaryocytes and platelets (Broudy et al., 1997; Fielder et al., 1996; 1997). However, multiple reports suggest that the regulation of serum TPO levels is much more complex and include IL-6 signaling (Burmester et al., 2005; Kaser et al., 2001; Wolber et al., 2001), clearance of aged platelets through Ashwell-Morell receptor (Grozovsky et al., 2015) and glycoprotein receptor GPIbα (Xu et al., 2018). On the translational level, steady-state serum TPO levels are maintained by the translational inhibition mechanism (Ghilardi et al., 1998). Disruption of this

regulation by different mutations within 5'-untranslated region causes hereditary thrombocythemia (Ghilardi and Skoda, 1999; Ghilardi et al., 1999; Kondo et al., 1998; Liu et al., 2008; Wiestner et al., 1998).

Here, we describe a novel activating mutation in the *THPO* gene leading to the hereditary thrombocytosis. G>T substitution resides in the 5'-UTR and disrupts the Kozak consensus sequence of uORF7, the main upstream open reading frame involved in the negative regulation of *THPO* mRNA translation. Mutations in the Kozak sequence were previously reported to be involved in the disease pathogenesis of β-thalassemia (De Angioletti et al., 2004), congenital heart defects (Mohan et al., 2014), androgen insensitivity syndrome (Choong et al., 1996), ataxia (Usuki, 2000), and breast cancer (Signori et al., 2001). We found that G>T transversion leads to the increase of TPO protein in the cell culture supernatant of HEK293 cells transfected with *THPO* G480T construct, in comparison to the *THPO* wildtype control. Furthermore, *in vitro* translation of the *THPO* G480T mRNA produced more TPO protein than the translation of *THPO* wildtype mRNA.

Our results demonstrate that this single nucleotide variant decreases the translation initiation rate from uORF7. This allows ribosomes to initiate the translation from AUG8, a start codon of TPO polypeptide. Interestingly, G480T mutation shows a weak phenotype compared to the previously reported mutations. This observation could be explained by the fact that the guanine at +4 position is only one of the required residues of the Kozak consensus sequence (Kozak, 2002).

3.3 Concluding remarks

In conclusion, we identified a novel mutation in the erythropoietin gene that causes an inherited form of erythrocytosis. We discovered an unexpected mode of action of this mutation, in which an alternative non-coding mRNAs originating in intron 1 of the *EPO* gene connect to the *EPO* open reading frame due to the frameshift inflicted by the mutation and produce the excessive quantities of EPO protein. Alternative transcripts originating in intron 1 of other genes may be involved in the pathogenesis of the diseases, as recently demonstrated for the *VHL* gene and erythrocytosis (Lenglet et al., 2018). Our finding has a direct clinical impact since the sequencing of the *EPO* gene should be implemented into the standard diagnostic pipeline for idiopathic erythrocytosis. Moreover, the results of our study suggest that intron 1 transcripts may regulate *EPO* expression in the liver, a process that is to date poorly

understood. Lastly, our study nicely demonstrates the usefulness and suitability of genome-editing techniques in identifying the molecular mechanisms of rare disorders.

We also described a novel gain-of-function variant in the thrombopoietin gene that gives rise to the familial thrombocytosis. This novel variant modifies the Kozak sequence of the upstream open reading frame, disrupts the negative regulation of *THPO* expression, and leads to the elevated serum TPO levels.

4 References

Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L.A.M., et al. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell *121*, 295–306.

Aird, W.C., Parvin, J.D., Sharp, P.A., and Rosenberg, R.D. (1994). The interaction of GATA-binding proteins and basal transcription factors with GATA box-containing core promoters. A model of tissue-specific gene expression. J. Biol. Chem. *269*, 883–889.

Akashi, K., Traver, D., Miyamoto, T., and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature *404*, 193–197.

Alexander, W.S. (2002). Suppressors of cytokine signalling (SOCS) in the immune system. Nat Rev Immunol *2*, 410–416.

Anderson, D.M., Anderson, K.M., Chang, C.-L., Makarewich, C.A., Nelson, B.R., McAnally, J.R., Kasaragod, P., Shelton, J.M., Liou, J., Bassel-Duby, R., et al. (2015). A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell *160*, 595–606.

Ang, S.O., Chen, H., Hirota, K., Gordeuk, V.R., Jelinek, J., Guan, Y., Liu, E., Sergueeva, A.I., Miasnikova, G.Y., Mole, D., et al. (2002). Disruption of oxygen homeostasis underlies congenital Chuvash polycythemia. Nat. Genet. *32*, 614–621.

Arber, D.A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M.J., Le Beau, M.M., Bloomfield, C.D., Cazzola, M., and Vardiman, J.W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood *127*, 2391–2405.

Arcasoy, M.O. (2008). The non-haematopoietic biological effects of erythropoietin. Br. J. Haematol. *141*, 14–31.

Bachmann, S., Le Hir, M., and Eckardt, K.U. (1993). 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. *41*, 335–341.

Baggiolini, A., Varum, S., Mateos, J.M., Bettosini, D., John, N., Bonalli, M., Ziegler, U., Dimou, L., Clevers, H., Furrer, R., et al. (2015). Premigratory and Migratory Neural Crest Cells Are Multipotent In Vivo. Stem Cell *16*, 314–322.

Ballmaier, M., Germeshausen, M., Schulze, H., Cherkaoui, K., Lang, S., Gaudig, A., Krukemeier, S., Eilers, M., Strauss, G., and Welte, K. (2001). c-mpl mutations are the cause of congenital amegakaryocytic thrombocytopenia. Blood *97*, 139–146.

Barrett, T.D., Palomino, H.L., Brondstetter, T.I., Kanelakis, K.C., Wu, X., Haug, P.V., Yan, W., Young, A., Hua, H., Hart, J.C., et al. (2011). Pharmacological Characterization of 1-(5-Chloro-6-(trifluoromethoxy)-1H-benzoimidazol-2-yl)-1H-pyrazole-4-carboxylic Acid (JNJ-42041935), a Potent and Selective Hypoxia-Inducible Factor Prolyl Hydroxylase Inhibitor. Molecular Pharmacology *79*, 910–920.

Bartley, T.D., Bogenberger, J., Hunt, P., Li, Y.S., Lu, H.S., Martin, F., Chang, M.S., Samal, B., Nichol, J.L., and Swift, S. (1994). Identification and cloning of a megakaryocyte growth and development factor that is a ligand for the cytokine receptor Mpl. Cell *77*, 1117–1124.

Bastian, L.S., Kwiatkowski, B.A., Breininger, J., Danner, S., and Roth, G. (1999). Regulation of the megakaryocytic glycoprotein IX promoter by the oncogenic Ets transcription factor Fli-1. Blood *93*, 2637–2644.

Basu, P., Lung, T.K., Lemsaddek, W., Sargent, T.G., Williams, D.C., Basu, M., Redmond, L.C., Lingrel, J.B., Haar, J.L., and Lloyd, J.A. (2007). EKLF and KLF2 have compensatory roles in embryonic beta-globin gene expression and primitive erythropoiesis. Blood *110*, 3417–3425.

Baxter, E.J., Scott, L.M., Campbell, P.J., East, C., Fourouclas, N., Swanton, S., Vassiliou, G.S., Bench, A.J., Boyd, E.M., Curtin, N., et al. (2005). Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. The Lancet *365*, 1054–1061.

Beaulieu, E., Green, L., Elsby, L., Alourfi, Z., Morand, E.F., Ray, D.W., and Donn, R. (2011). Identification of a novel cell type-specific intronic enhancer of macrophage migration inhibitory factor (MIF) and its regulation by mithramycin. Clin. Exp. Immunol. *163*, 178–188.

Bellanne-Chantelot, C., Chaumarel, I., Labopin, M., Bellanger, F., Barbu, V., De Toma, C., Delhommeau, F., Casadevall, N., Vainchenker, W., Thomas, G., et al. (2006). Genetic and clinical implications of the Val617Phe JAK2 mutation in 72 families with myeloproliferative disorders. Blood *108*, 346–352.

Benveniste, P., Frelin, C., Janmohamed, S., Barbara, M., Herrington, R., Hyam, D., and Iscove, N.N. (2010). Intermediate-Term Hematopoietic Stem Cells with Extended but Time-Limited Reconstitution Potential. Stem Cell *6*, 48–58.

Bernaudin, M., Bellail, A., Marti, H.H., Yvon, A., Vivien, D., Duchatelle, I., MacKenzie, E.T., and Petit, E. (2000). Neurons and astrocytes express EPO mRNA: Oxygen-sensing mechanisms that involve the redox-state of the brain. Glia *30*, 271–278.

Bernhardt, W.M., Wiesener, M.S., Scigalla, P., Chou, J., Schmieder, R.E., Gunzler, V., and Eckardt, K.U. (2010). Inhibition of Prolyl Hydroxylases Increases Erythropoietin Production in ESRD. Journal of the American Society of Nephrology *21*, 2151–2156.

Bertoncello, I., Hodgson, G.S., and Bradley, T.R. (1985). Multiparameter analysis of transplantable hemopoietic stem cells: I. The separation and enrichment of stem cells homing to marrow and spleen on the basis of rhodamine-123 fluorescence. Exp. Hematol. *13*, 999–1006.

Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y.R., and Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. Nature *464*, 108–111.

Besarab, A., Bolton, W.K., Browne, J.K., Egrie, J.C., Nissenson, A.R., Okamoto, D.M., Schwab, S.J., and Goodkin, D.A. (1998). The effects of normal as compared with low hematocrit values in patients with cardiac disease who are receiving hemodialysis and epoetin. N. Engl. J. Med. *339*, 584–590.

Bianchi, M., Crinelli, R., Giacomini, E., Carloni, E., and Magnani, M. (2009). A potent enhancer element in the 5'-UTR intron is crucial for transcriptional regulation of the human ubiquitin C gene. Gene *448*, 88–101.

Blanchard, K.L., Acquaviva, A.M., Galson, D.L., and Bunn, H.F. (1992). Hypoxic induction of the human erythropoietin gene: cooperation between the promoter and enhancer, each of which contains steroid receptor response elements. Mol. Cell. Biol. *12*, 5373–5385.

Bodo, E., Kromminga, A., Funk, W., Laugsch, M., Duske, U., Jelkmann, W., and Paus, R. (2007). Human hair follicles are an extrarenal source and a nonhematopoietic target of erythropoietin. Faseb J. *21*, 3346–3354.

Boettcher, M., Lentini, S., Arens, E.R., Kaiser, A., van der Mey, D., Thuss, U., Kubitza, D., and Wensing, G. (2018). First-in-man-proof of concept study with molidustat: a novel selective oral HIF-prolyl hydroxylase inhibitor for the treatment of renal anaemia. British Journal of Clinical Pharmacology *84*, 1557–1565.

Bondurant, M.C., Koury, M.J., Koury, S.T., and Semenza, G. (1991). Erythropoietin ontogeny and organ distribution in mice. Semin. Hematol. *28*, 20–5–discussion26–7.

Boudville, N.C., Djurdjev, O., Macdougall, I.C., de Francisco, A.L.M., Deray, G., Besarab, A., Stevens, P.E., Walker, R.G., Urena, P., Inigo, P., et al. (2009). Hemoglobin Variability in Nondialysis Chronic Kidney Disease: Examining the Association with Mortality. Clin J Am Soc Nephrol *4*, 1176–1182.

Briddell, R.A., Brandt, J.E., Straneva, J.E., Srour, E.F., and Hoffman, R. (1989). Characterization of the Human Burst-Forming Unit-Megakaryocyte. Blood *74*, 145–151.

Brigandi, R.A., Ariazi, J.L., Duffy, K.J., Luo, L., Nephrol, D.A.J.A.S., 2010 The prolyl-hydroxylase inhibitor, GSK1278863A, induced EPO in vitro and efficient erythropoiesis leading to increased hemoglobin in vivo.

Broudy, V.C., Lin, N.L., Sabath, D.F., Papayannopoulou, T., and Kaushansky, K. (1997). Human platelets display high-affinity receptors for thrombopoietin. Blood *89*, 1896–1904.

Bruick, R.K., and McKnight, S.L. (2001). A conserved family of prolyl-4-hydroxylases that modify HIF. Science *294*, 1337–1340.

Bunn, H.F. (2013). Erythropoietin. Cold Spring Harbor Perspectives in Medicine *3*, a011619–a011619.

Burmester, H., Wolber, E.-M., Freitag, P., Fandrey, J., and Jelkmann, W. (2005). Thrombopoietin production in wild-type and interleukin-6 knockout mice with acute inflammation. J. Interferon Cytokine Res. *25*, 407–413.

Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S.M., Reth, M., Höfer, T., and Rodewald, H.-R. (2015). Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. Nature *518*, 542–546.

Bussel, J.B., Kuter, D.J., George, J.N., McMillan, R., Aledort, L.M., Conklin, G.T., Lichtin, A.E., Lyons, R.M., Nieva, J., Wasser, J.S., et al. (2006). AMG 531, a thrombopoiesis-stimulating protein, for chronic ITP. N. Engl. J. Med. *355*, 1672–1681.

Cario, H., Schwarz, K., Jorch, N., Kyank, U., Petrides, P.E., Schneider, D.T., Uhle, R., Debatin, K.M., and Kohne, E. (2005). Mutations in the von Hippel-Lindau (VHL) tumor suppressor gene and VHL-haplotype analysis in patients with presumable congenital erythrocytosis. Haematologica *90*, 19–24.

Carnot, P., and Deflandre, C. (1906). The hemopoietic activity of different organs during the regeneration of blood. Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences *143*, 432–435.

Cazzola, M., and Skoda, R.C. (2000). Translational pathophysiology: a novel molecular mechanism of human disease. Blood *95*, 3280–3288.

Charache, S., Weatherall, D.J., and Clegg, J.B. (1966). Polycythemia associated with a hemoglobinopathy. J. Clin. Invest. *45*, 813–822.

Chasis, J.A., and Mohandas, N. (2008). Erythroblastic islands: niches for erythropoiesis. Blood *112*, 470–478.

Chen, S., Su, Y., and Wang, J. (2013). ROS-mediated platelet generation: a microenvironment-dependent manner for megakaryocyte proliferation, differentiation and maturation. Cell Death & Disease *4*, e722–e722.

Choi, E.S., Nichol, J.L., Hokom, M.M., Hornkohl, A.C., and Hunt, P. (1995). Platelets generated in vitro from proplatelet-displaying human megakaryocytes are functional. Blood *85*, 402–413.

Choong, C.S., Quigley, C.A., French, F.S., and Wilson, E.M. (1996). A novel missense mutation in the amino-terminal domain of the human androgen receptor gene in a family with partial androgen insensitivity syndrome causes reduced efficiency of protein translation. J. Clin. Invest. *98*, 1423–1431.

Coşkun, S., Chao, H., Vasavada, H., Heydari, K., Gonzales, N., Zhou, X., de Crombrugghe, B., and Hirschi, K.K. (2014). Development of the fetal bone marrow niche and regulation of HSC quiescence and homing ability by emerging osteolineage cells. CellReports *9*, 581–590.

Coulon, V., Chebli, K., Cavelier, P., and Blanchard, J.-M. (2010). A Novel Mouse c-fos Intronic Promoter That Responds to CREB and AP-1 Is Developmentally Regulated In Vivo. PLoS ONE *5*, e11235–11.

Cramer, E.M., Norol, F., Guichard, J., BretonGorius, J., Vainchenker, W., Masse, J.M., and Debili, N. (1997). Ultrastructure of platelet formation by human megakaryocytes cultured with the Mpl ligand. Blood *89*, 2336–2346.

Crane, G.M., Jeffery, E., and Morrison, S.J. (2017). Adult haematopoietic stem cell niches. Nat Rev Immunol *17*, 573–590.

Crispino, J.D. (2005). GATA1 in normal and malignant hematopoiesis. Semin. Cell Dev. Biol. *16*, 137–147.

Cwirla, S.E., Balasubramanian, P., Duffin, D.J., Wagstrom, C.R., Gates, C.M., Singer, S.C., Davis, A.M., Tansik, R.L., Mattheakis, L.C., Boytos, C.M., et al. (1997). Peptide agonist of the thrombopoietin receptor as potent as the natural cytokine. Science *276*, 1696–1699.

Dame, C., Fahnenstich, H., Freitag, P., Hofmann, D., Abdul-Nour, T., Bartmann, P., and Fandrey, J. (1998). Erythropoietin mRNA expression in human fetal and neonatal tissue. Blood *92*, 3218–3225.

Dame, C., Sola, M.C., Lim, K.-C., Leach, K.M., Fandrey, J., Ma, Y., Knöpfle, G., Engel, J.D., and Bungert, J. (2004). Hepatic Erythropoietin Gene Regulation by GATA-4. J. Biol. Chem. *279*, 2955–2961.

Darnell, J.E. (1997). STATs and gene regulation. Science *277*, 1630–1635.

Darnell, J., Kerr, I., and Stark, G. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science *264*, 1415–1421.

De Angioletti, M., Lacerra, G., Sabato, V., and Carestia, C. (2004). β+45 G → C: A novel silent β-thalassaemia mutation, the first in the Kozak sequence. Br. J. Haematol. *124*, 224–231.

De Maria, R., Zeuner, A., Eramo, A., Domenichelli, C., Bonci, D., Grignani, F., Srinivasula, S.M., Alnemri, E.S., Testa, U., and Peschle, C. (1999). Negative regulation of erythropoiesis by caspase-mediated cleavage of GATA-1. Nature *401*, 489–493.

de Sauvage, F.J. (1996). Physiological regulation of early and late stages of megakaryocytopoiesis by thrombopoietin. The Journal of Experimental Medicine *183*, 651–656.

de Sauvage, F.J., Hass, P.E., Spencer, S.D., Malloy, B.E., Gurney, A.L., Spencer, S.A., Darbonne, W.C., Henzel, W.J., Wong, S.C., and Kuang, W.J. (1994). Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl ligand. Nature *369*, 533–538.

de Seigneux, S., Lundby, A.K.M., Berchtold, L., Berg, A.H., Saudan, P., and Lundby, C. (2016). Increased Synthesis of Liver Erythropoietin with CKD. Journal of the American Society of Nephrology *27*, 2265–2269.

Decker, M., Leslie, J., Liu, Q., and Ding, L. (2018). Hepatic thrombopoietin is required for bone marrow hematopoietic stem cell maintenance. Science *360*, 106–110.

Deveaux, S., CohenKaminsky, S., Shivdasani, R.A., Andrews, N.C., Filipe, A., Kuzniak, I., Orkin, S.H., Roméo, P.H., and Mignotte, V. (1997). p45 NF-E2 regulates expression of thromboxane synthase in megakaryocytes. The EMBO Journal *16*, 5654–5661.

Ding, J. (2004). Familial essential thrombocythemia associated with a dominant-positive activating mutation of the c-MPL gene, which encodes for the receptor for thrombopoietin. Blood *103*, 4198–4200.

Drueke, T.B., Locatelli, F., Clyne, N., Eckardt, K.-U., Macdougall, I.C., Tsakiris, D., Burger, H.-U., Scherhag, A., and Investigators, C. (2006). Normalization of hemoglobin level in patients with chronic kidney disease and anemia. N. Engl. J. Med. *355*, 2071–2084.

Durbin, J.E., Hackenmiller, R., Simon, M.C., and Levy, D.E. (1996). Targeted disruption of the mouse Stat1 gene results in compromised innate immunity to viral disease. Cell *84*, 443–450.

Dzierzak, E., and Philipsen, S. (2013). Erythropoiesis: development and differentiation. Cold Spring Harbor Perspectives in Medicine *3*, a011601–a011601.

Eckardt, K.U., Ratcliffe, P.J., Tan, C.C., Bauer, C., and Kurtz, A. (1992). Age-dependent expression of the erythropoietin gene in rat liver and kidneys. J. Clin. Invest. *89*, 753–760.

Eilken, H.M., Nishikawa, S.-I., and Schroeder, T. (2009). Continuous single-cell imaging of blood generation from haemogenic endothelium. Nature *457*, 896–900.

El-Harith, E.-H.A., Roesl, C., Ballmaier, M., Germeshausen, M., Frye-Boukhriss, H., Neuhoff, von, N., Becker, C., Nuernberg, G., Nuernberg, P., Ahmed, M.A.M., et al. (2009). Familial thrombocytosis caused by the novel germ-line mutation p.Pro106Leu in the MPL gene. Br. J. Haematol. *144*, 185–194.

Elagib, K.E., Racke, F.K., Mogass, M., Khetawat, R., Delehanty, L.L., and Goldfarb, A.N. (2003). RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic differentiation. Blood *101*, 4333–4341.

Endo, T.A., Masuhara, M., Yokouchi, M., Suzuki, R., Sakamoto, H., Mitsui, K., Matsumoto, A., Tanimura, S., Ohtsubo, M., Misawa, H., et al. (1997). A new protein containing an SH2 domain that inhibits JAK kinases. Nature *387*, 921–924.

Epstein, A.C., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A., et al. (2001). C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. Cell *107*, 43–54.

Erickson-Miller, C., Delorme, E., Giampa, L., Hopson, C., Valoret, E., Tian, S.S., Miller, S.G., Keenan, R., Rosen, J., Dillon, S., et al. (2004). Biological activity and selectivity for Tpo receptor of the orally bioavailable, small molecule Tpo receptor agonist, SB-497115. Blood *104*, 796A–796A.

Erslev, A. (1953). Humoral Regulation of Red Cell Production. Blood *8*, 349–357.

Erslev, A.J. (1991). Drug-Therapy - Erythropoietin. N. Engl. J. Med. *324*, 1339–1344.

Falke, L.L., Gholizadeh, S., Goldschmeding, R., Kok, R.J., and Nguyen, T.Q. (2015). Diverse origins of the myofibroblast -implications for kidney fibrosis. Nat Rev Nephrol *11*, 233–244.

Fandrey, J., and Bunn, H.F. (1993). In vivo and in vitro regulation of erythropoietin mRNA: measurement by competitive polymerase chain reaction. Blood *81*, 617–623.

Fielder, P.J., Gurney, A.L., Stefanich, E., Marian, M., Moore, M.W., Carver-Moore, K., and de Sauvage, F.J. (1996). Regulation of thrombopoietin levels by c-mpl-mediated binding to platelets. Blood *87*, 2154–2161.

Fielder, P.J., Hass, P., Nagel, M., Stefanich, E., Widmer, R., Bennett, G.L., Keller, G.A., de Sauvage, F.J., and Eaton, D. (1997). Human platelets as a model for the binding and degradation of thrombopoietin. Blood *89*, 2782–2788.

Foster, D.C., Sprecher, C.A., Grant, F.J., Kramer, J.M., Kuijper, J.L., Holly, R.D., Whitmore, T.E., Heipel, M.D., Bell, L.A., and Ching, A.F. (1994). Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosomal localization. Pnas *91*, 13023–13027.

Franco, R.S. (2012). Measurement of red cell lifespan and aging. Transfus Med Hemother *39*, 302–307.

Fried, W. (1972). The liver as a source of extrarenal erythropoietin production. Blood *40*, 671–677.

Fried, W., Kilbridge, T., Krantz, S., McDonald, T.P., and Lange, R.D. (1969). Studies on Extrarenal Erythropoietin. J. Lab. Clin. Med. *73*, 244–.

Fu, X.-Y., and Zhang, J.-J. (1993). Transcription factor p91 interacts with the epidermal growth factor receptor and mediates activation of the c-fos gene promoter. Cell *74*, 1135–1145.

Fujiwara, Y., Browne, C.P., Cunniff, K., Goff, S.C., and Orkin, S.H. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. Proc. Natl. Acad. Sci. U.S.a. *93*, 12355–12358.

Gaunitz, F., Deichsel, D., Heise, K., Werth, M., Anderegg, U., and Gebhardt, R. (2005). An intronic silencer element is responsible for specific zonal expression of glutamine synthetase in the rat liver. Hepatology *41*, 1225–1232.

Gaunitz, F., Heise, K., and Gebhardt, R. (2004). A Silencer Element in the First Intron of the Glutamine Synthetase Gene Represses Induction by Glucocorticoids. Molecular Endocrinology *18*, 63–69.

Gekas, C., Dieterlen-Lièvre, F., Orkin, S.H., and Mikkola, H.K.A. (2005). The Placenta Is a Niche for Hematopoietic Stem Cells. Developmental Cell *8*, 365–375.

Ghilardi, N., and Skoda, R.C. (1999). A single-base deletion in the thrombopoietin (TPO) gene causes familial essential thrombocythemia through a mechanism of more efficient translation of TPO mRNA. Blood *94*, 1480–1482.

Ghilardi, N., Wiestner, A., and Skoda, R.C. (1998). Thrombopoietin production is inhibited by a translational mechanism. Blood *92*, 4023–4030.

Ghilardi, N., Wiestner, A., Kikuchi, M., Ohsaka, A., and Skoda, R.C. (1999). Hereditary thrombocythaemia in a Japanese family is caused by a novel point mutation in the thrombopoietin gene. Br. J. Haematol. *107*, 310–316.

Goldberg, M.A., Glass, G.A., Cunningham, J.M., and Bunn, H.F. (1987). The regulated expression of erythropoietin by two human hepatoma cell lines. Proc. Natl. Acad. Sci. U.S.a. *84*, 7972–7976.

Goldwasser, E., Kung, C., and Eliason, J. (1974). Mechanism of Erythropoietin-Induced Differentiation .13. Role of Sialic-Acid in Erythropoietin Action. J. Biol. Chem. *249*, 4202–4206.

Goodell, M.A. (1996). Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. The Journal of Experimental Medicine *183*, 1797–1806.

Graziano, C., Carone, S., Panza, E., Marino, F., Magini, P., Romeo, G., Pession, A., and Seri, M. (2009). Association of hereditary thrombocythemia and distal limb defects with a thrombopoietin gene mutation. Blood *114*, 1655–1657.

Gregoli, P.A., and Bondurant, M.C. (1999). Function of caspases in regulating apoptosis caused by erythropoietin deprivation in erythroid progenitors. J. Cell. Physiol. *178*, 133–143.

Gregory, C.J., and Eaves, A.C. (1977). Human marrow cells capable of erythropoietic differentiation in vitro: definition of three erythroid colony responses. Blood *49*, 855–864.

Griesshammer, M., Hornkohl, A., Nichol, J.L., Hecht, T., Raghavachar, A., Heimpel, H., and Schrezenmeier, H. (1998). High levels of thrombopoietin in sera of patients with essential thrombocythemia: cause or consequence of abnormal platelet production? Ann. Hematol. *77*, 211–215.

Gross, A.W., and Lodish, H.F. (2006). Cellular trafficking and degradation of erythropoietin and novel erythropoiesis stimulating protein (NESP). J. Biol. Chem. *281*, 2024–2032.

Grote, D., Souabni, A., Busslinger, M., and Bouchard, M. (2006). Pax2/8-regulated Gata3 expression is necessary for morphogenesis and guidance of the nephric duct in the developing kidney. Development *133*, 53–61.

Grover, A., Sanjuan-Pla, A., Thongjuea, S., Carrelha, J., Giustacchini, A., Gambardella, A., Macaulay, I., Mancini, E., Luis, T.C., Mead, A., et al. (2016). Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. Nat Commun *7*, 11075.

Grozovsky, R., Begonja, A.J., Liu, K., Visner, G., Hartwig, J.H., Falet, H., and Hoffmeister, K.M. (2015). The Ashwell-Morell receptor regulates hepatic thrombopoietin production via JAK2-STAT3 signaling. Nature Medicine *21*, 47–54.

Gruber, M., Hu, C.-J., Johnson, R.S., Brown, E.J., Keith, B., and Simon, M.C. (2007). Acute postnatal ablation of Hif-2alpha results in anemia. Pnas *104*, 2301–2306.

Gu, Y.Z., Moran, S.M., Hogenesch, J.B., Wartman, L., and Bradfield, C.A. (1998). Molecular characterization and chromosomal localization of a third alpha-class hypoxia inducible factor subunit, HIF3alpha. Gene Expr. *7*, 205–213.

Gurney, A.L., Carver-Moore, K., de Sauvage, F.J., and Moore, M.W. (1994). Thrombocytopenia in c-mpl-deficient mice. Science *265*, 1445–1447.

Haas, S., Hansson, J., Klimmeck, D., Loeffler, D., Velten, L., Uckelmann, H., Wurzer, S., Prendergast, Á.M., Schnell, A., Hexel, K., et al. (2015). Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. Stem Cell *17*, 1–14.

Haetscher, N., Feuermann, Y., Wingert, S., Rehage, M., Thalheimer, F.B., Weiser, C., Bohnenberger, H., Jung, K., Schroeder, T., Serve, H., et al. (2015). STAT5-regulated microRNA-193b controls haematopoietic stem and progenitor cell expansion by modulating cytokine receptor signalling. Nat Commun *6*, 1–11.

Halupa, A., Bailey, M.L., Huang, K., Iscove, N.N., Levy, D.E., and Barber, D.L. (2005). A novel role for STAT1 in regulating murine erythropoiesis: deletion of STAT1 results in overall reduction of erythroid progenitors and alters their distribution. Blood *105*, 552–561.

Hammaren, H.M., Ungureanu, D., Grisouard, J., Skoda, R.C., Hubbard, S.R., and Silvennoinen, O. (2015). ATP binding to the pseudokinase domain of JAK2 is critical for pathogenic activation. Pnas *112*, 4642–4647.

Hilton, D. (2001). SOCS proteins - negative regulators of the cytokine signal transduction. Biochemical Society Transactions *29*, A105.1–A105.

Hodge, D., Coghill, E., Keys, J., Maguire, T., Hartmann, B., McDowall, A., Weiss, M., Grimmond, S., and Perkins, A. (2006). A global role for EKLF in definitive and primitive erythropoiesis. Blood *107*, 3359–3370.

Hon, W.-C., Wilson, M.I., Harlos, K., Claridge, T.D.W., Schofield, C.J., Pugh, C.W., Maxwell, P.H., Ratcliffe, P.J., Stuart, D.I., and Jones, E.Y. (2002). Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. Nature *417*, 975–978.

Horvath, C.M., Wen, Z., and Darnell, J.E. (1995). A STAT protein domain that determines DNA sequence recognition suggests a novel DNA-binding domain. Genes & Development *9*, 984–994.

Hoyer, J.D., Allen, S.L., Beutler, E., Kubik, K., West, C., and Fairbanks, V.F. (2004). Erythrocytosis due to bisphosphoglycerate mutase deficiency with concurrent glucose-6-phosphate dehydrogenase (G-6-PD) deficiency. Am. J. Hematol. *75*, 205–208.

Huang, L.E., Arany, Z., Livingston, D.M., and Bunn, H.F. (1996). Activation of Hypoxia-inducible Transcription Factor Depends Primarily upon Redox-sensitive Stabilization of Its α Subunit. J. Biol. Chem. *271*, 32253–32259.

Huang, Z., Richmond, T.D., Muntean, A.G., Barber, D.L., Weiss, M.J., and Crispino, J.D. (2007). STAT1 promotes megakaryopoiesis downstream of GATA-1 in mice. J. Clin. Invest. *117*, 3890–3899.

Humphreys, B.D., Lin, S.-L., Kobayashi, A., Hudson, T.E., Nowlin, B.T., Bonventre, J.V., Valerius, M.T., McMahon, A.P., and Duffield, J.S. (2010). Fate tracing reveals the pericyte and not epithelial origin of myofibroblasts in kidney fibrosis. Am. J. Pathol. *176*, 85–97.

Hussein, K., Percy, M., and McMullin, M.F. (2012). Clinical utility gene card for: familial erythrocytosis. European Journal of Human Genetics *20*, 1–4.

Hussein, K., Percy, M., McMullin, M.F., Schwarz, J., Schnittger, S., Porret, N., Martinez-Aviles, L.M., Paricio, B.B., Giraudier, S., Skoda, R., et al. (2014). Clinical utility gene card for: hereditary thrombocythemia. European Journal of Human Genetics *22*, 1–5.

Ichikawa, N., Ishida, F., Shimodaira, S., Tahara, T., Kato, T., and Kitano, K. (1996). Regulation of serum thrombopoietin levels by platelets and megakaryocytes in patients with aplastic anaemia and idiopathic thrombocytopenic purpura. Thromb. Haemost. *76*, 156–160.

Ihara, K., Ishii, E., Eguchi, M., Takada, H., Suminoe, A., Good, R.A., and Hara, T. (1999). Identification of mutations in the c-mpl gene in congenital amegakaryocytic thrombocytopenia. Pnas *96*, 3132–3136.

Imagawa, S., Goldberg, M.A., Doweiko, J., and Bunn, H.F. (1991). Regulatory Elements of the Erythropoietin Gene. Blood *77*, 278–285.

Imagawa, S., Yamamoto, M., and Miura, Y. (1997). Negative regulation of the erythropoietin gene expression by the GATA transcription factors. Blood *89*, 1430–1439.

Imagawa, S., Suzuki, N., Ohmine, K., Obara, N., Mukai, H.Y., Ozawa, K., Yamamoto, M., and Nagasawa, T. (2002). GATA suppresses erythropoietin gene expression through GATA site in mouse erythropoietin gene promoter. Int J Hematol *75*, 376–381.

Italiano, J.E., Lecine, P., Shivdasani, R.A., and Hartwig, J.H. (1999). Blood platelets are assembled principally at the ends of proplatelet processes produced by differentiated megakaryocytes. J. Cell Biol. *147*, 1299–1312.

Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J.M., Lane, W.S., and Kaelin, W.G. (2001). HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. Science *292*, 464–468.

Ivanovs, A., Rybtsov, S., Welch, L., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2011). Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-gonad-mesonephros region. The Journal of Experimental Medicine *208*, 2417–2427.

Jaakkola, P., Mole, D.R., Tian, Y.M., Wilson, M.I., Gielbert, J., Gaskell, S.J., Kriegsheim, von, A., Hebestreit, H.F., Mukherji, M., Schofield, C.J., et al. (2001). Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. Science *292*, 468–472.

Jacobson, L.O., Goldwasser, E., Fried, W., and Plzak, L. (1957). Role of the Kidney in Erythropoiesis. Nature *179*, 633–634.

James, C., Ugo, V., Le Couédic, J.-P., Staerk, J., Delhommeau, F., Lacout, C., Garçon, L., Raslova, H., Berger, R., Bennaceur-Griscelli, A., et al. (2005). A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. Nature *434*, 1144–1148.

Jelkmann, W. (2007). Erythropoietin after a century of research: younger than ever. Eur. J. Haematol. *78*, 183–205.

Jo, D., Lin, Q., Nashabi, A., Mays, D.J., Unutmaz, D., Pietenpol, J.A., and Ruley, H.E. (2003). Cell cycle-dependent transduction of cell-permeant Cre recombinase proteins. J. Cell. Biochem. *89*, 674–687.

Johnstone, R.M. (1992). The Jeanne Manery-Fisher Memorial Lecture 1991. Maturation of reticulocytes: formation of exosomes as a mechanism for shedding membrane proteins. Biochem. Cell Biol. *70*, 179–190.

Junt, T., Schulze, H., Chen, Z., Massberg, S., Goerge, T., Krueger, A., Wagner, D.D., Graf, T., Italiano, J.E., Shivdasani, R.A., et al. (2007). Dynamic Visualization of Thrombopoiesis Within Bone Marrow. Science *317*, 1767–1770.

Kaneko, K., Furuyama, K., Fujiwara, T., Kobayashi, R., Ishida, H., Harigae, H., and Shibahara, S. (2014). Identification of a novel erythroid-specific enhancer for the ALAS2 gene and its loss-of-function mutation which is associated with congenital sideroblastic anemia. Haematologica *99*, 252–261.

Kapitsinou, P.P., Liu, Q., Unger, T.L., Rha, J., Davidoff, O., Keith, B., Epstein, J.A., Moores, S.L., Erickson-Miller, C.L., and Haase, V.H. (2010). Hepatic HIF-2 regulates erythropoietic responses to hypoxia in renal anemia. Blood *116*, 3039–3048.

Kapralova, K., Horvathova, M., Pecquet, C., Kucerova, J.F., Pospisilova, D., Leroy, E., Kralova, B., Feenstra, J.D.M., Schischlik, F., Kralovics, R., et al. (2016). Cooperation of germ line JAK2 mutations E846D and R1063H in hereditary erythrocytosis with megakaryocytic atypia. Blood *128*, 1418–1423.

Kaser, A., Brandacher, G., Steurer, W., Kaser, S., Offner, F.A., Zoller, H., Theurl, I., Widder, W., Molnar, C., Ludwiczek, O., et al. (2001). Interleukin-6 stimulates thrombopoiesis through thrombopoietin: role in inflammatory thrombocytosis. Blood *98*, 2720–2725.

Kaushansky, K. (2005). The molecular mechanisms that control thrombopoiesis. J. Clin. Invest. *115*, 3339–3347.

Kaushansky, K. (2006). Mechanisms of disease: Lineage-specific hematopoietic growth factors. N. Engl. J. Med. *354*, 2034–2045.

Kaushansky, K., Broudy, V.C., Lin, N., Jorgensen, M.J., McCarty, J., Fox, N., Zucker- Franklin, D., and Lofton-Day, C. (1995). Thrombopoietin, the Mpl ligand, is essential for full megakaryocyte development. Proc. Natl. Acad. Sci. U.S.a. *92*, 3234–3238.

Kaushansky, K., Lichtman, M.A., Prchal, J.T., and Levi, M. (2016). Williams Hematology, 9th edition (McGraw-Hill Education).

Kaushansky, K., Lok, S., Holly, R.D., Broudy, V.C., Lin, N., Bailey, M.C., Forstrom, J.W., Buddle, M.M., Oort, P.J., and Hagen, F.S. (1994). Promotion of megakaryocyte progenitor expansion and differentiation by the c-Mpl ligand thrombopoietin. Nature *369*, 568–571.

Kelemen, E., Cserháti, I., and Tanos, B. (1958). Demonstration and some properties of human thrombopoietin in thrombocythaemic sera. Acta Haematol. *20*, 350–355.

Kelley, L.L., Koury, M.J., Bondurant, M.C., Koury, S.T., Sawyer, S.T., and Wickrema, A. (1993). Survival or death of individual proerythroblasts results from differing erythropoietin sensitivities: a mechanism for controlled rates of erythrocyte production. Blood *82*, 2340–2352.

Keohane, C., McMullin, M.F., and Harrison, C. (2013). The diagnosis and management of erythrocytosis. Bmj *347*, f6667–f6667.

Kiel, M.J., Yilmaz, O.H., Iwashita, T., Yilmaz, O.H., Terhorst, C., and Morrison, S.J. (2005). SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. Cell *121*, 1109–1121.

Kikuchi, M., Tayama, T., Hayakawa, H., Takahashi, I., Hoshino, H., and Ohsaka, A. (1995). Familial thrombocytosis. Br. J. Haematol. *89*, 900–902.

Kim, A.R., Ulirsch, J.C., Wilmes, S., Unal, E., Moraga, I., Karakukcu, M., Yuan, D., Kazerounian, S., Abdulhay, N.J., King, D.S., et al. (2017). Functional Selectivity in Cytokine Signaling Revealed Through a Pathogenic EPO Mutation. Cell *168*, 1053–1064.e15.

Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A.K., Nagy, A., and Tam, P.P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. Development *126*, 4691–4701.

Kissa, K., and Herbomel, P. (2010). Blood stem cells emerge from aortic endothelium by a novel type of cell transition. Nature *464*, 112–115.

Klampfl, T., Gisslinger, H., Harutyunyan, A.S., Nivarthi, H., Rumi, E., Milosevic, J.D., Them, N.C.C., Berg, T., Gisslinger, B., Pietra, D., et al. (2013). Somatic Mutations of Calreticulin in Myeloproliferative Neoplasms. N. Engl. J. Med. *369*, 2379–2390.

Kondo, M., Weissman, I.L., and Akashi, K. (1997). Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. Cell *91*, 661–672.

Kondo, T., Okabe, M., Sanada, M., Kurosawa, M., Suzuki, S., Kobayashi, M., Hosokawa, M., and Asaka, M. (1998). Familial essential thrombocythemia associated with one-base deletion in the 5 '-untranslated region of the thrombopoietin gene. Blood *92*, 1091–1096.

Kondo, T., Plaza, S., Zanet, J., Benrabah, E., Valenti, P., Hashimoto, Y., Kobayashi, S., Payre, F., and Kageyama, Y. (2010). Small peptides switch the transcriptional activity of Shavenbaby during Drosophila embryogenesis. Science *329*, 336–339.

Kosugi, S., Kurata, Y., Tomiyama, Y., Tahara, T., Kato, T., Tadokoro, S., Shiraga, M., Honda, S., Kanakura, Y., and Matsuzawa, Y. (1996). Circulating thrombopoietin level in chronic immune thrombocytopenic purpura. Br. J. Haematol. *93*, 704–706.

Koury, M.J., and Bondurant, M.C. (1990a). Erythropoietin retards DNA breakdown and prevents programmed death in erythroid progenitor cells. Science *248*, 378–381.

Koury, M.J., and Bondurant, M.C. (1990b). Erythropoietin retards DNA breakdown and prevents programmed death in erythroid progenitor cells. Science *248*, 378–381.

Koury, M.J., and Haase, V.H. (2015). Anaemia in kidney disease: harnessing hypoxia responses for therapy. Nat Rev Nephrol *11*, 394–410.

Koury, S.T., Bondurant, M.C., and Koury, M.J. (1988). Localization of erythropoietin synthesizing cells in murine kidneys by in situ hybridization. Blood *71*, 524–527.

Koury, S.T., Bondurant, M.C., Koury, M.J., and Semenza, G.L. (1991). Localization of cells producing erythropoietin in murine liver by in situ hybridization. Blood *77*, 2497–2503.

Koury, S.T., Koury, M.J., Bondurant, M.C., Caro, J., and Graber, S.E. (1989). Quantitation of erythropoietin-producing cells in kidneys of mice by in situ hybridization: correlation with hematocrit, renal erythropoietin mRNA, and serum erythropoietin concentration. Blood *74*, 645–651.

Kozak, M. (2002). Emerging links between initiation of translation and human diseases. Mamm. Genome *13*, 401–410.

Kralovics, R., Passamonti, F., Buser, A.S., Teo, S.-S., Tiedt, R., Passweg, J.R., Tichelli, A., Cazzola, M., and Skoda, R.C. (2005). A gain-of-function mutation of JAK2 in myeloproliferative disorders. N. Engl. J. Med. *352*, 1779–1790.

Kunisaki, Y., Bruns, I., Scheiermann, C., Ahmed, J., Pinho, S., Zhang, D., Mizoguchi, T., Wei, Q., Lucas, D., Ito, K., et al. (2013). Arteriolar niches maintain haematopoietic stem cell quiescence. Nature *502*, 637–.

Kuo, C.T., Morrisey, E.E., Anandappa, R., Sigrist, K., Lu, M.M., Parmacek, M.S., Soudais, C., and Leiden, J.M. (1997). GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. Genes & Development *11*, 1048–1060.

Kuter, D.J., and Rosenberg, R.D. (1995). The reciprocal relationship of thrombopoietin (c-Mpl ligand) to changes in the platelet mass during busulfan-induced thrombocytopenia in the rabbit. Blood *85*, 2720–2730.

La Chapelle, De, A., Träskelin, A.L., and Juvonen, E. (1993). Truncated erythropoietin receptor causes dominantly inherited benign human erythrocytosis. Proc. Natl. Acad. Sci. U.S.a. *90*, 4495–4499.

Lacombe, C., Da Silva, J.L., Bruneval, P., Fournier, J.G., Wendling, F., Casadevall, N., Camilleri, J.P., Bariety, J., Varet, B., and Tambourin, P. (1988). Peritubular cells are the site of erythropoietin synthesis in the murine hypoxic kidney. J. Clin. Invest. *81*, 620–623.

Ladroue, C., Carcenac, R., Leporrier, M., Gad, S., Le Hello, C., Galateau-Salle, F., Feunteun, J., Pouysségur, J., Richard, S., and Gardie, B. (2008). PHD2 mutation and congenital erythrocytosis with paraganglioma. N. Engl. J. Med. *359*, 2685–2692.

Landolfi, R., Marchioli, R., Kutti, J., Gisslinger, H., Tognoni, G., Patrono, C., Barbui, T., European Collaboration on Low-Dose Aspirin in Polycythemia Vera Investigators (2004). Efficacy and safety of low-dose aspirin in polycythemia vera. N. Engl. J. Med. *350*, 114–124.

Lawson, K.A., Meneses, J.J., and Pedersen, R.A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. Development *113*, 891–911.

Lecine, P., Italiano, J.E., Kim, S.W., Villeval, J.L., and Shivdasani, R.A. (2000). Hematopoietic-specific beta 1 tubulin participates in a pathway of platelet biogenesis dependent on the transcription factor NF-E2. Blood *96*, 1366–1373.

Lecine, P., Villeval, J.L., Vyas, P., Swencki, B., Xu, Y.H., and Shivdasani, R.A. (1998). Mice lacking transcription factor NF-E2 provide in vivo validation of the proplatelet model of thrombocytopoiesis and show a platelet production defect that is intrinsic to megakaryocytes. Blood *92*, 1608–1616.

Lee, P.J., Jiang, B.H., Chin, B.Y., Iyer, N.V., Alam, J., Semenza, G.L., and Choi, A. (1997). Hypoxia-inducible factor-1 mediates transcriptional activation of the heme oxygenase-1 gene in response to hypoxia. J. Biol. Chem. *272*, 5375–5381.

Lenglet, M., Robriquet, F., Schwarz, K., Camps, C., Couturier, A., Hoogewijs, D., Buffet, A., Knight, S.J.L., Gad, S., Couve, S., et al. (2018). Identification of a new VHL exon and complex splicing alterations in familial erythrocytosis or von Hippel-Lindau disease. Blood *132*, 469–483.

Lentjes, M.H.F.M., Niessen, H.E.C., Akiyama, Y., de Bruïne, A.P., Melotte, V., and van Engeland, M. (2016). The emerging role of GATA transcription factors in development and disease. Expert Rev Mol Med *18*, e3.

Levine, R.L., Wadleigh, M., Cools, J., Ebert, B.L., Wernig, G., Huntly, B.J.P., Boggon, T.J., Wlodarska, I., Clark, J.J., Moore, S., et al. (2005). Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell *7*, 387–397.

Levy, D.E., and Darnell, J.E. (2002). STATs: transcriptional control and biological impact. Nature Reviews Molecular Cell Biology *3*, 651–662.

Li, J.Z., Yang, C., Xia, Y.P., Bertino, A., Glaspy, J., Roberts, M., and Kuter, D.J. (2001). Thrombocytopenia caused by the development of antibodies to thrombopoietin. Blood *98*, 3241–3248.

Lin, S.-L., Kisseleva, T., Brenner, D.A., and Duffield, J.S. (2008). Pericytes and Perivascular Fibroblasts Are the Primary Source of Collagen-Producing Cells in Obstructive Fibrosis of the Kidney. Am. J. Pathol. *173*, 1617–1627.

Liu, K., Kralovics, R., Rudzki, Z., Grabowska, B., Buser, A.S., Olcaydu, D., Gisslinger, H., Tiedt, R., Frank, P., Okoñ, K., et al. (2008). A de novo splice donor mutation in the thrombopoietin gene causes hereditary thrombocythemia in a Polish family. Haematologica *93*, 706–714.

Liu, K., Martini, M., Rocca, B., Amos, C.I., Teofili, L., Giona, F., Ding, J., Komatsu, H., Larocca, L.M., and Skoda, R.C. (2009). Evidence for a founder effect of the MPL-S505N mutation in eight Italian pedigrees with hereditary thrombocythemia. Haematologica *94*, 1368–1374.

Liu, Y., Pop, R., Sadegh, C., Brugnara, C., Haase, V.H., and Socolovsky, M. (2006). Suppression of Fas-FasL coexpression by erythropoietin mediates erythroblast expansion during the erythropoietic stress response in vivo. Blood *108*, 123–133.

Livnah, O., Stura, E.A., Middleton, S.A., Johnson, D.L., Jolliffe, L.K., and Wilson, I.A. (1999). Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. Science *283*, 987–990.

Lok, C.N., and Ponka, P. (1999). Identification of a hypoxia response element in the transferrin receptor gene. J. Biol. Chem. *274*, 24147–24152.

Lok, S., Kaushansky, K., Holly, R.D., Kuijper, J.L., Lofton-Day, C.E., Oort, P.J., Grant, F.J., Heipel, M.D., Burkhead, S.K., Kramer, J.M., et al. (1994). Cloning and expression of murine thrombopoietin cDNA and stimulation of platelet production in vivo. Nature *369*, 565–568.

Long, M.W. (1998). Megakaryocyte differentiation events. Semin. Hematol. *35*, 192–199.

Long, M.W., Gragowski, L.L., Heffner, C.H., and Boxer, L.A. (1985). Phorbol diesters stimulate the development of an early murine progenitor cell. The burst-forming unit-megakaryocyte. J. Clin. Invest. *76*, 431–438.

Loya, F., Yang, Y., Lin, H., Goldwasser, E., and Albitar, M. (1994). Transgenic Mice Carrying the Erythropoietin Gene Promoter Linked to Lacz Express the Reporter in Proximal Convoluted Tubule Cells After Hypoxia. Blood *84*, 1831–1836.

Lönnberg, M., Andrén, M., Birgegård, G., Drevin, M., Garle, M., and Carlsson, J. (2012). Rapid detection of erythropoiesis-stimulating agents in urine and serum. Analytical Biochemistry *420*, 101–114.

Lundberg, P., Takizawa, H., Kubovcakova, L., Guo, G., Hao-Shen, H., Dirnhofer, S., Orkin, S.H., Manz, M.G., and Skoda, R.C. (2014). Myeloproliferative neoplasms can be initiated from a single hematopoietic stem cell expressing JAK2-V617F. Journal of Experimental Medicine *211*, 2213–2230.

Madan, A., Lin, C., Hatch, S.L., and Curtin, P.T. (1995). Regulated basal, inducible, and tissue-specific human erythropoietin gene expression in transgenic mice requires multiple cis DNA sequences. Blood *85*, 2735–2741.

Magny, E.G., Pueyo, J.I., Pearl, F.M.G., Cespedes, M.A., Niven, J.E., Bishop, S.A., and Couso, J.P. (2013). Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. Science *341*, 1116–1120.

Majeti, R., Park, C.Y., and Weissman, I.L. (2007). Identification of a Hierarchy of Multipotent Hematopoietic Progenitors in Human Cord Blood. Cell Stem Cell *1*, 635–645.

Marti, H.H., Wenger, R.H., Rivas, L.A., Straumann, U., Digicaylioglu, M., Henn, V., Yonekawa, Y., Bauer, C., and Gassmann, M. (1996). Erythropoietin gene expression in human, monkey and murine brain. European Journal of Neuroscience *8*, 666–676.

Maslah, N., Cassinat, B., Verger, E., Kiladjian, J.-J., and Velazquez, L. (2017). The role of LNK/SH2B3 genetic alterations in myeloproliferative neoplasms and other hematological disorders. Leukemia *31*, 1661–1670.

Mastrogiannaki, M., Matak, P., Keith, B., Simon, M.C., Vaulont, S., and Peyssonnaux, C. (2009). HIF-2alpha, but not HIF-1alpha, promotes iron absorption in mice. J. Clin. Invest. *119*, 1159–1166.

Maxwell, A.P., Lappin, T.R.J., Johnston, C.F., Bridges, J.M., and McGeown, M.G. (1990). Erythropoietin production in kidney tubular cells. Br. J. Haematol. *74*, 535–539.

Maxwell, P.H., Ferguson, D.J., Osmond, M.K., Pugh, C.W., Heryet, A., Doe, B.G., Johnson, M.H., and Ratcliffe, P.J. (1994). Expression of a homologously recombined erythopoietin-SV40 T antigen fusion gene in mouse liver: evidence for erythropoietin production by Ito cells. Blood *84*, 1823–1830.

McLeod, D.L., Shreve, M.M., and Axelrad, A.A. (1976). Induction of megakaryocyte colonies with platelet formation in vitro. Nature *261*, 492–494.

Mead, A.J., Rugless, M.J., Jacobsen, S.E.W., and Schuh, A. (2012). Germline JAK2 Mutation in a Family with Hereditary Thrombocytosis. N. Engl. J. Med. *366*, 967–969.

Medvinsky, A., and Dzierzak, E. (1996). Definitive Hematopoiesis Is Autonomously Initiated by the AGM Region. Cell *86*, 897–906.

Mehta, J., Wang, H., Iqbal, S.U., and Mesa, R. (2014). Epidemiology of myeloproliferative neoplasms in the United States. Leuk. Lymphoma *55*, 595–600.

Merryweather-Clarke, A.T., Atzberger, A., Soneji, S., Gray, N., Clark, K., Waugh, C., McGowan, S.J., Taylor, S., Nandi, A.K., Wood, W.G., et al. (2011). Global gene expression analysis of human erythroid progenitors. Blood *117*, e96–e108.

Meyer, S.C., Keller, M.D., Woods, B.A., LaFave, L.M., Bastian, L., Kleppe, M., Bhagwat, N., Marubayashi, S., and Levine, R.L. (2014). Genetic studies reveal an unexpected negative regulatory role for Jak2 in thrombopoiesis. Blood *124*, 2280–2284.

Mignotte, V., Vigon, I., Boucher de Crèvecoeur, E., Roméo, P.H., Lemarchandel, V., and Chrétien, S. (1994). Structure and transcription of the human c-mpl gene (MPL). Genomics *20*, 5–12.

Miller, I.J., and Bieker, J.J. (1993). A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. Mol. Cell. Biol. *13*, 2776–2786.

Minamishima, Y.A., and Kaelin, W.G. (2010). Reactivation of Hepatic EPO Synthesis in Mice After PHD Loss. Science *329*, 407–407.

Mirand, E.A., Murphy, G.P., Steeves, R.A., Groenewald, J.M., and Deklerk, J.N. (1969). Erythropoietin Activity in Anephric Allotransplanted Unilaterally Nephrectomized and Intact Man. J. Lab. Clin. Med. *73*, 121–.

Mirand, E.A., Murphy, G.P., Steeves, R.A., Weber, H.W., and Retief, F.P. (1968). Extra-renal production of erythropoietin in man. Acta Haematol. *39*, 359–365.

Miro-Murillo, M., Elorza, A., Soro-Arnaiz, I., Albacete-Albacete, L., Ordonez, A., Balsa, E., Vara-Vega, A., Vazquez, S., Fuertes, E., Fernandez-Criado, C., et al. (2011). Acute Vhl Gene Inactivation Induces Cardiac HIF-Dependent Erythropoietin Gene Expression. PLoS ONE *6*.

Miyakawa, Y., Oda, A., Druker, B.J., Miyazaki, H., Handa, M., Ohashi, H., and Ikeda, Y. (1996). Thrombopoietin induces tyrosine phosphorylation of Stat3 and Stat5 in human blood platelets. Blood *87*, 439–446.

Miyake, T., Kung, C., and Goldwasser, E. (1977). Purification of Human Erythropoietin. J. Biol. Chem. *252*, 5558–5564.

Mohan, R.A., van Engelen, K., Stefanovic, S., Barnett, P., Ilgun, A., Baars, M.J.H., Bouma, B.J., Mulder, B.J.M., Christoffels, V.M., and Postma, A.V. (2014). A mutation in the Kozak sequence of GATA4 hampers translation in a family with atrial septal defects. Am. J. Med. Genet. *164*, 2732–2738.

Moliterno, A.R., Williams, D.M., Gutierrez-Alamillo, L.I., Salvatori, R., Ingersoll, R.G., and Spivak, J.L. (2004). Mpl Baltimore: A thrombopoietin receptor polymorphism associated with thrombocytosis. Pnas *101*, 11444–11447.

Morrison, S.J., and Weissman, I.L. (1994). The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity *1*, 661–673.

Mujais, S.K., Beru, N., Pullman, T.N., and Goldwasser, E. (1999). Erythropoietin is produced by tubular cells of the rat kidney. Cell Biochem. Biophys. *30*, 153–166.

Muta, K., and Krantz, S.B. (1993). Apoptosis of human erythroid colony-forming cells is decreased by stem cell factor and insulin-like growth factor I as well as erythropoietin. J. Cell. Physiol. *156*, 264–271.

Müller, A.M., Medvinsky, A., Strouboulis, J., Grosveld, F., and Dzierzak, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. Immunity *1*, 291–301.

Nagata, Y., Muro, Y., and Todokoro, K. (1997). Thrombopoietin-induced polyploidization of bone marrow megakaryocytes is due to a unique regulatory mechanism in late mitosis. J. Cell Biol. *139*, 449–457.

Naka, T., Narazaki, M., Hirata, M., Matsumoto, T., Minamoto, S., Aono, A., Nishimoto, N., Kajita, T., Taga, T., Yoshizaki, K., et al. (1997). Structure and function of a new STAT-induced STAT inhibitor. Nature *387*, 924–929.

Nakamura-Ishizu, A., Takizawa, H., and Suda, T. (2014). The analysis, roles and regulation of quiescence in hematopoietic stem cells. Development *141*, 4656–4666.

Nangalia, J., Massie, C.E., Baxter, E.J., Nice, F.L., Gundem, G., Wedge, D.C., Avezov, E., Li, J., Kollmann, K., Kent, D.G., et al. (2013). Somatic CALR Mutations in Myeloproliferative Neoplasms with Nonmutated JAK2. N. Engl. J. Med. *369*, 2391–2405.

Nathan, D.G., Schupak, E., Stohlman, F., and Merrill, J.P. (1964). Erythropoiesis in Anephric Man. J. Clin. Invest. *43*, 2158–.

Nelson, B.R., Makarewich, C.A., Anderson, D.M., Winders, B.R., Troupes, C.D., Wu, F., Reese, A.L., McAnally, J.R., Chen, X., Kavalali, E.T., et al. (2016). A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science *351*, 271–275.

Neubauer, H., Cumano, A., Müller, M., Wu, H., Huffstadt, U., and Pfeffer, K. (1998). Jak2 deficiency defines an essential developmental checkpoint in definitive hematopoiesis. Cell *93*, 397–409.

Ney, P.A. (2011). Normal and disordered reticulocyte maturation. Curr. Opin. Hematol. *18*, 152–157.

Ng, A.P., Kauppi, M., Metcalf, D., Hyland, C.D., Josefsson, E.C., Lebois, M., Zhang, J.-G., Baldwin, T.M., Di Rago, L., Hilton, D.J., et al. (2014). Mpl expression on megakaryocytes and platelets is dispensable for thrombopoiesis but essential to prevent myeloproliferation. Proc. Natl. Acad. Sci. U.S.a. *111*, 5884–5889.

Nichols, K.E., Crispino, J.D., Poncz, M., White, J.G., Orkin, S.H., Maris, J.M., and Weiss, M.J. (2000). Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. Nat. Genet. *24*, 266–270.

Nimmo, R.A., May, G.E., and Enver, T. (2015). Primed and ready: understanding lineage commitment through single cell analysis. Trends in Cell Biology *25*, 459–467.

Noguchi, C.T. (2008). Where the Epo cells are. Blood *111*, 4836–4837.

Nombela-Arrieta, C., Pivarnik, G., Winkel, B., Canty, K.J., Harley, B., Mahoney, J.E., Park, S.-Y., Lu, J., Protopopov, A., and Silberstein, L.E. (2013). Quantitative imaging of haematopoietic stem and progenitor cell localization and hypoxic status in the bone marrow microenvironment. Nat. Cell Biol. *15*, 533–543.

Norkin, M., and Wingard, J.R. (2017). Recent advances in hematopoietic stem cell transplantation. F1000Res *6*, 870.

Notta, F., Doulatov, S., Laurenti, E., Poeppl, A., Jurisica, I., and Dick, J.E. (2011). Isolation of Single Human Hematopoietic Stem Cells Capable of Long-Term Multilineage Engraftment. Science *333*, 218–221.

Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O.I., Wilson, G., Kaufmann, K.B., McLeod, J., Laurenti, E., Dunant, C.F., et al. (2016). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science *351*, aab2116–aab2116.

Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R., and Grosveld, F. (1995). Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature *375*, 316–318.

Obara, N., Suzuki, N., Kim, K., Nagasawa, T., Imagawa, S., and Yamamoto, M. (2008). Repression via the GATA box is essential for tissue-specific erythropoietin gene expression. Blood *111*, 5223–5232.

Okada, S., Nakauchi, H., Nagayoshi, K., Nishikawa, S., Miura, Y., and Suda, T. (1992). In vivo and in vitro stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. Blood *80*, 3044–3050.

Osawa, M., Hanada, K.I., Hamada, H., and Nakauchi, H. (1996). Long-Term Lymphohematopoietic Reconstitution by a Single CD34-Low/Negative Hematopoietic Stem Cell. Science *273*, 242–245.

Ott, C.J., Suszko, M., Blackledge, N.P., Wright, J.E., Crawford, G.E., and Harris, A. (2009). A complex intronic enhancer regulates expression of the CFTR gene by direct interaction with the promoter. Journal of Cellular and Molecular Medicine *13*, 680–692.

Ottersbach, K., and Dzierzak, E. (2005). The Murine Placenta Contains Hematopoietic Stem Cells within the Vascular Labyrinth Region. Developmental Cell *8*, 377–387.

Pan, X., Suzuki, N., Hirano, I., Yamazaki, S., Minegishi, N., and Yamamoto, M. (2011). Isolation and characterization of renal erythropoietin-producing cells from genetically produced anemia mice. PLoS ONE *6*, e25839–11.

Pandolfi, P.P., Roth, M.E., Karis, A., Leonard, M.W., Dzierzak, E., Grosveld, F.G., Engel, J.D., and Lindenbaum, M.H. (1995). Targeted Disruption of the Gata3 Gene Causes Severe Abnormalities in the Nervous-System and in Fetal Liver Hematopoiesis. Nat. Genet. *11*, 40–44.

Pardanani, A.D., Levine, R.L., Lasho, T., Pikman, Y., Mesa, R.A., Wadleigh, M., Steensma, D.P., Elliott, M.A., Wolanskyj, A.P., Hogan, W.J., et al. (2006). MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. Blood *108*, 3472–3476.

Parganas, E., Wang, D., Stravopodis, D., Topham, D.J., Marine, J.C., Teglund, S., Vanin, E.F., Bodner, S., Colamonici, O.R., van Deursen, J.M., et al. (1998). Jak2 is essential for signaling through a variety of cytokine receptors. Cell *93*, 385–395.

Parkins, A.C., Sharpe, A.H., and Orkin, S.H. (1995). Lethal β-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature *375*, 318–322.

Passegué, E., Wagers, A.J., Giuriato, S., Anderson, W.C., and Weissman, I.L. (2005). Global analysis of proliferation and cell cycle gene expression in the regulation of hematopoietic stem and progenitor cell fates. The Journal of Experimental Medicine *202*, 1599–1611.

Passweg, J.R., Baldomero, H., Bader, P., Bonini, C., Cesaro, S., Dreger, P., Duarte, R.F., Dufour, C., Kuball, J., Farge-Bancel, D., et al. (2016). Hematopoietic stem cell transplantation in Europe 2014: more than 40 000 transplants annually. Bone Marrow Transplantation *51*, 786–792.

Pauli, A., Norris, M.L., Valen, E., Chew, G.-L., Gagnon, J.A., Zimmerman, S., Mitchell, A., Ma, J., Dubrulle, J., Reyon, D., et al. (2014). Toddler: an embryonic signal that promotes cell movement via Apelin receptors. Science *343*, 1248636–1248636.

Pearson, M.A., Reczek, D., Bretscher, A., and Karplus, P.A. (2000). Structure of the ERM protein moesin reveals the FERM domain fold masked by an extended actin binding tail domain. Cell *101*, 259–270.

Pecci, A., Ragab, I., Bozzi, V., De Rocco, D., Barozzi, S., Giangregorio, T., Ali, H., Melazzini, F., Sallam, M., Alfano, C., et al. (2018). Thrombopoietin mutation in congenital amegakaryocytic thrombocytopenia treatable with romiplostim. EMBO Mol Med *10*, 63–75.

Percy, M.J., Furlow, P.W., Lucas, G.S., Li, X., Lappin, T.R.J., McMullin, M.F., and Lee, F.S. (2008). A gain-of-function mutation in the HIF2A gene in familial erythrocytosis. N. Engl. J. Med. *358*, 162–168.

Perié, L., Duffy, K.R., Kok, L., de Boer, R.J., and Schumacher, T.N. (2015). The Branching Point in Erythro-Myeloid Differentiation. Cell *163*, 1655–1662.

Peschle, C., Marone, G., Genovese, A., Cillo, C., Magli, C., and Condorelli, M. (1975). Erythropoietin production by the liver in fetal-neonatal life. Life Sci. *17*, 1325–1330.

Pevny, L., Lin, C.S., D'Agati, V., Simon, M.C., Orkin, S.H., and Costantini, F. (1995). Development of hematopoietic cells lacking transcription factor GATA-1. Development *121*, 163–172.

Pfeffer, M.A., Burdmann, E.A., Chen, C.-Y., Cooper, M.E., de Zeeuw, D., Eckardt, K.-U., Feyzi, J.M., Ivanovich, P., Kewalramani, R., Levey, A.S., et al. (2009). A Trial of Darbepoetin Alfa in Type 2 Diabetes and Chronic Kidney Disease. N. Engl. J. Med. *361*, 2019–2032.

Pikman, Y., Lee, B.H., Mercher, T., McDowell, E., Ebert, B.L., Gozo, M., Cuker, A., Wernig, G., Moore, S., Galinsky, I., et al. (2006). MPLW515L Is a Novel Somatic Activating Mutation in Myelofibrosis with Myeloid Metaplasia. PLoS Medicine *3*, e270.

Pope, N.J., and Bresnick, E.H. (2010). Differential coregulator requirements for function of the hematopoietic transcription factor GATA-1 at endogenous loci. Nucleic Acids Res *38*, 2190–2200.

Provenzano, R., Besarab, A., Sun, C.H., Diamond, S.A., Durham, J.H., Cangiano, J.L., Aiello, J.R., Novak, J.E., Lee, T., Leong, R., et al. (2016). Oral Hypoxia-Inducible Factor Prolyl Hydroxylase Inhibitor Roxadustat (FG-4592) for the Treatment of Anemia in Patients with CKD. Clin J Am Soc Nephrol *11*, 982–991.

Qian, H., Buza-Vidas, N., Hyland, C.D., Jensen, C.T., Antonchuk, J., Månsson, R., Thoren, L.A., Ekblom, M., Alexander, W.S., and Jacobsen, S.E.W. (2007). Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. Cell Stem Cell *1*, 671–684.

Quintás-Cardama, A., Kantarjian, H., Cortes, J., and Verstovsek, S. (2011). Janus kinase inhibitors for the treatment of myeloproliferative neoplasias and beyond. Nat Rev Drug Discov *10*, 127–140.

Radley, J.M., and Haller, C.J. (1982). The Demarcation Membrane System of the Megakaryocyte - a Misnomer. Blood *60*, 213–219.

Rankin, E.B., Biju, M.P., Liu, Q., Unger, T.L., Rha, J., Johnson, R.S., Simon, M.C., Keith, B., and Haase, V.H. (2007). Hypoxia-inducible factor-2 (HIF-2) regulates hepatic erythropoietin in vivo. J. Clin. Invest. *117*, 1068–1077.

Rankin, E.B., Wu, C., Khatri, R., Wilson, T.L.S., Andersen, R., Araldi, E., Rankin, A.L., Yuan, J., Kuo, C.J., Schipani, E., et al. (2012). The HIF Signaling Pathway in Osteoblasts Directly Modulates Erythropoiesis through the Production of EPO. Cell *149*, 63–74.

Rearick, D., Prakash, A., McSweeny, A., Shepard, S.S., Fedorova, L., and Fedorov, A. (2011). Critical association of ncRNA with introns. Nucleic Acids Res *39*, 2357–2366.

Reinert, R.B., Kantz, J., Misfeldt, A.A., Poffenberger, G., Gannon, M., Brissova, M., and Powers, A.C. (2012). Tamoxifen-Induced Cre-loxP Recombination Is Prolonged in Pancreatic Islets of Adult Mice. PLoS ONE *7*, e33529.

Richmond, T.D., Chohan, M., and Barber, D.L. (2005). Turning cells red: signal transduction mediated by erythropoietin. Trends in Cell Biology *15*, 146–155.

Rieger, M.A., and Schroeder, T. (2012). Hematopoiesis. Cold Spring Harbor Perspectives in Biology *4*, a008250–a008250.

Rodriguez-Fraticelli, A.E., Wolock, S.L., Weinreb, C.S., Panero, R., Patel, S.H., Jankovic, M., Sun, J., Calogero, R.A., Klein, A.M., and Camargo, F.D. (2018). Clonal analysis of lineage fate in native haematopoiesis. Nature *553*, 212–216.

Rolfs, A., Kvietikova, I., Gassmann, M., and Wenger, R.H. (1997). Oxygen-regulated transferrin expression is mediated by hypoxia-inducible factor-1. J. Biol. Chem. *272*, 20055–20062.

Rongvaux, A., Willinger, T., Martinek, J., Strowig, T., Gearty, S.V., Teichmann, L.L., Saito, Y., Marches, F., Halene, S., Palucka, A.K., et al. (2014). Development and function of human innate immune cells in a humanized mouse model. Nat Biotechnol *32*, 364–U230.

Rubiolo, C., Piazzolla, D., Meissl, K., Beug, H., Huber, J.C., Kolbus, A., and Baccarini, M. (2006). A balance between Raf-1 and Fas expression sets the pace of erythroid differentiation. Blood *108*, 152–159.

Rumi, E., Harutyunyan, A.S., Pietra, D., Feenstra, J.D.M., Cavalloni, C., Roncoroni, E., Casetti, I., Bellini, M., Milanesi, C., Renna, M.C., et al. (2016). LNK mutations in familial myeloproliferative neoplasms. Blood *128*, 144–145.

Saleh, M.N., Bussel, J.B., Cheng, G., Meyer, O., Bailey, C.K., Arning, M., Brainsky, A., on behalf of the EXTEND Study Group (2013). Safety and efficacy of eltrombopag for treatment of chronic immune thrombocytopenia: results of the long-term, open-label EXTEND study. Blood *121*, 537–545.

Sanjuan-Pla, A., Macaulay, I.C., Jensen, C.T., Woll, P.S., Luis, T.C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T., et al. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature Publishing Group *502*, 232–236.

Sankaran, V.G., Xu, J., and Orkin, S.H. (2010). Advances in the understanding of haemoglobin switching. Br. J. Haematol. *149*, 181–194.

Sasaki, A., Yasukawa, H., Shouda, T., Kitamura, T., Dikic, I., and Yoshimura, A. (2000). CIS3/SOCS-3 suppresses erythropoietin (EPO) signaling by binding the EPO receptor and JAK2. J. Biol. Chem. *275*, 29338–29347.

Schepers, A.G., Snippert, H.J., Stange, D.E., van den Born, M., van Es, J.H., van de Wetering, M., and Clevers, H. (2012). Lineage Tracing Reveals Lgr5+ Stem Cell Activity in Mouse Intestinal Adenomas. Science *337*, 730–735.

Schlemper, R.J., van der Maas, A.P.C., and Eikenboom, J.C.J. (1994). Familial essential thrombocythemia: Clinical characteristics of 11 cases in one family. Ann. Hematol. *68*, 153–158.

Schödel, J., Oikonomopoulos, S., Ragoussis, J., Pugh, C.W., Ratcliffe, P.J., and Mole, D.R. (2011). High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. Blood *117*, e207–e217.

Scohy, S., Gabant, P., Szpirer, C., and Szpirer, J. (2000). Identification of an enhancer and an alternative promoter in the first intron of the alpha-fetoprotein gene. Nucleic Acids Res *28*, 3743–3751.

Scortegagna, M., Ding, K., Zhang, Q., Oktay, Y., Bennett, M.J., Bennett, M., Shelton, J.M., Richardson, J.A., Moe, O., and Garcia, J.A. (2005). HIF-2alpha regulates murine hematopoietic development in an erythropoietin-dependent manner. Blood *105*, 3133–3140.

Scortegagna, M., Morris, M.A., Oktay, Y., Bennett, M., and Garcia, J.A. (2003). The HIF family member EPAS1/HIF-2alpha is required for normal hematopoiesis in mice. Blood *102*, 1634–1640.

Semenza, G.L., Koury, S.T., NEJFELT, M.K., GEARHART, J.D., and ANTONARAKIS, S.E. (1991). Cell-Type-Specific and Hypoxia-Inducible Expression of the Human Erythropoietin Gene in Transgenic Mice. Pnas *88*, 8725–8729.

Semenza, G.L., Dureza, R.C., Traystman, M.D., Gearhart, J.D., and Antonarakis, S.E. (1990). Human erythropoietin gene expression in transgenic mice: multiple transcription initiation sites and cis-acting regulatory elements. Mol. Cell. Biol. *10*, 930–938.

Sender, R., Fuchs, S., and Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. PLoS Biol *14*, e1002533–14.

Shah, Y.M., Matsubara, T., Ito, S., Yim, S.-H., and Gonzalez, F.J. (2009). Intestinal Hypoxia-inducible Transcription Factors Are Essential for Iron Absorption following Iron Deficiency. Cell Metab. *9*, 152–164.

Shimizu, R., Takahashi, S., Ohneda, K., Engel, J.D., and Yamamoto, M. (2001). In vivo requirements for GATA-1 functional domains during primitive and definitive erythropoiesis. The EMBO Journal *20*, 5250–5260.

Shimoda, K., Feng, J., Murakami, H., Nagata, S., Watling, D., Rogers, N.C., Stark, G.R., Kerr, I.M., and Ihle, J.N. (1997). Jak1 plays an essential role for receptor phosphorylation and Stat activation in response to granulocyte colony-stimulating factor. Blood *90*, 597–604.

Shinjo, K., Takeshita, A., Nakamura, S., Naitoh, K., Yanagi, M., Tobita, T., Ohnishi, K., and Ohno, R. (1998). Serum thrombopoietin levels in patients correlate inversely with platelet counts during chemotherapy-induced thrombocytopenia. Leukemia *12*, 295–300.

Shivdasani, R.A., Fujiwara, Y., McDevitt, M.A., and Orkin, S.H. (1997). A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. The EMBO Journal *16*, 3965–3973.

Shivdasani, R.A., Rosenblatt, M.F., Zucker-Franklin, D., Jackson, C.W., Hunt, P., Saris, C.J.M., and Orkin, S.H. (1995). Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoeitin/MGDF in megakaryocyte development. Cell *81*, 695–704.

Shoemaker, C.B., and Mitsock, L.D. (1986). Murine erythropoietin gene: cloning, expression, and human gene homology. Mol. Cell. Biol. *6*, 849–858.

Shuai, K. (1994). Interferon activation of the transcription factor Stat91 involves dimerization through SH2-phosphotyrosyl peptide interactions. Cell *76*, 821–828.

Shuai, K., Stark, G.R., Kerr, I.M., and Darnell, J.E. (1993). A single phosphotyrosine residue of Stat91 required for gene activation by interferon-gamma. Science *261*, 1744–1746.

Shuai, K., and Liu, B. (2003). Regulation of JAK–STAT signalling in the immune system. Nat Rev Immunol *3*, 900–911.

Signori, E., Bagni, C., Papa, S., Primerano, B., Rinaldi, M., Amaldi, F., and Fazio, V.M. (2001). A somatic mutation in the 5′UTR of BRCA1 gene in sporadic breast cancer causes down-modulation of translation efficiency. Oncogene *20*, 4596–4600.

Simon, M.C. (1995). Gotta have GATA. Nat. Genet. *11*, 9–11.

Singh, A.K., Szczech, L., Tang, K.L., Barnhart, H., Sapp, S., Wolfson, M., Reddan, D., and Investigators, C. (2006). Correction of anemia with epoetin alfa in chronic kidney disease. N. Engl. J. Med. *355*, 2085–2098.

Skoda, R.C., Seldin, D.C., Chiang, M.K., Peichel, C.L., Vogt, T.F., and Leder, P. (1993). Murine c-mpl: a member of the hematopoietic growth factor receptor superfamily that transduces a proliferative signal. The EMBO Journal *12*, 2645–2653.

Snippert, H.J., van der Flier, L.G., Sato, T., van Es, J.H., van den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., van Rheenen, J., Simons, B.D., et al. (2010). Intestinal Crypt Homeostasis Results from Neutral Competition between Symmetrically Dividing Lgr5 Stem Cells. Cell *143*, 134–144.

Socolovsky, M., Fallon, A.E.J., Wang, S., Brugnara, C., and Lodish, H.F. (1999). Fetal Anemia and Apoptosis of Red Cell Progenitors in Stat5a−/−5b−/− Mice. Cell *98*, 181–191.

Sokol, L., Luhovy, M., Guan, Y.L., Prchal, J.F., Semenza, G.L., and Prchal, J.T. (1995). Primary Familial Polycythemia - a Frameshift Mutation in the Erythropoietin Receptor Gene and Increased Sensitivity of Erythroid Progenitors to Erythropoietin. Blood *86*, 15–22.

Song, W.J., Sullivan, M.G., Legare, R.D., Hutchings, S., Tan, X.L., Kufrin, D., Ratajczak, J., Resende, I.C., Haworth, C., Hock, R., et al. (1999). Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. Nat. Genet. *23*, 166–175.

Souma, T., Yamazaki, S., Moriguchi, T., Suzuki, N., Hirano, I., Pan, X., Minegishi, N., Abe, M., Kiyomoto, H., Ito, S., et al. (2013). Plasticity of Renal Erythropoietin-Producing Cells Governs Fibrosis. Journal of the American Society of Nephrology *24*, 1599–1616.

Southwood, C.M., Downs, K.M., and Bieker, J.J. (1996). Erythroid Krüppel-like factor exhibits an early and sequentially localized pattern of expression during mammalian erythroid ontogeny. Dev. Dyn. *206*, 248–259.

Souyri, M., Vigon, I., Penciolelli, J.F., Heard, J.M., Tambourin, P., and Wendling, F. (1990). A putative truncated cytokine receptor gene transduced by the myeloproliferative leukemia virus immortalizes hematopoietic progenitors. Cell *63*, 1137–1147.

Spangrude, G., Heimfeld, S., and Weissman, I. (1988). Purification and characterization of mouse hematopoietic stem cells. Science *241*, 58–62.

St Laurent, G., Shtokalo, D., Tackett, M.R., Yang, Z., Eremina, T., Wahlestedt, C., Urcuqui-Inchima, S., Seilheimer, B., McCaffrey, T.A., and Kapranov, P. (2012). Intronic RNAs constitute the major fraction of the non-coding RNA in mammalian cells. BMC Genomics *13*, 504.

Stark, G.R., Kerr, I.M., Williams, B., Silverman, R.H., and Schreiber, R.D. (1998). How cells respond to interferons. Annu. Rev. Biochem. *67*, 227–264.

Starr, R., Willson, T.A., Viney, E.M., Murray, L., Rayner, J.R., Jenkins, B.J., Gonda, T.J., Alexander, W.S., Metcalf, D., Nicola, N.A., et al. (1997). A family of cytokine-inducible inhibitors of signalling. Nature *387*, 917–921.

Stockklausner, C., Echner, N., Klotter, A.-C., Hegenbart, U., Dreger, P., and Kulozik, A.E. (2012). Hereditary thrombocythemia caused by a thrombopoietin (THPO) gain-of-function mutation associated with multiple myeloma and congenital limb defects. Ann. Hematol. *91*, 1129–1133.

Stoffel, R., Wiestner, A., and Skoda, R.C. (1995). Thrombopoietin in thrombocytopenic mice: Evidence against regulation at the mRNA level and for a direct regulatory role of platelets. Blood *86*, 3626–3626.

Storti, F., Santambrogio, S., Crowther, L.M., Otto, T., Abreu-Rodríguez, I., Kaufmann, M., Hu, C.-J., Dame, C., Fandrey, J., Wenger, R.H., et al. (2014). A novel distal upstream hypoxia response element regulating oxygen-dependent erythropoietin gene expression. Haematologica *99*, e45–e48.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J.B., Le, L., Ho, Y.-J., Klein, A., Hofmann, O., and Camargo, F.D. (2014). Clonal dynamics of native haematopoiesis. Nature *514*, 322–327.

Sungaran, R., Markovic, B., and Chong, B.H. (1997). Localization and regulation of thrombopoietin mRNA expression in human kidney, liver, bone marrow, and spleen using in situ hybridization. Blood *89*, 101–107.

Suzuki, N., Obara, N., Pan, X., Watanabe, M., Jishage, K.I., Minegishi, N., and Yamamoto, M. (2011). Specific Contribution of the Erythropoietin Gene 3' Enhancer to Hepatic Erythropoiesis after Late Embryonic Stages. Mol. Cell. Biol. *31*, 3896–3905.

Takeda, K., Noguchi, K., Shi, W., Tanaka, T., Matsumoto, M., Yoshida, N., Kishimoto, T., and Akira, S. (1997). Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. Pnas *94*, 3801–3804.

Takeda, K., Kaisho, T., Yoshida, N., Takeda, J., Kishimoto, T., and Akira, S. (1998). Stat3 activation is responsible for IL-6-dependent T cell proliferation through preventing apoptosis: Generation and characterization of T cell- specific stat3-deficient mice. J. Immunol. *161*, 4652–4660.

Tavian, M., and Peault, B. (2005). Embryonic development of the human hematopoietic system. The International Journal of Developmental Biology *49*, 243–250.

Taylor, M., Qu, A., Anderson, E.R., Matsubara, T., Martin, A., Gonzalez, F.J., and Shah, Y.M. (2011). Hypoxia-inducible factor-2α mediates the adaptive increase of intestinal ferroportin during iron deficiency in mice. Gastroenterology *140*, 2044–2055.

Teglund, S., McKay, C., Schuetz, E., van Deursen, J.M., Stravopodis, D., Wang, D., Brown, M., Bodner, S., Grosveld, G., and Ihle, J.N. (1998). Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. Cell *93*, 841–850.

Teofili, L., and Larocca, L.M. (2011). Advances in understanding the pathogenesis of familial thrombocythaemia. Br. J. Haematol. *152*, 701–712.

Teofili, L., Giona, F., Torti, L., Cenci, T., Ricerca, B.M., Rumi, C., Nunes, V., Foà, R., Leone, G., Martini, M., et al. (2010). Hereditary thrombocytosis caused by MPLSer505Asn is associated with a high thrombotic risk, splenomegaly and progression to bone marrow fibrosis. Haematologica *95*, 65–70.

Thomas, E.D., Lochte, H.L., Jr., Lu, W.C., and Ferrebee, J.W. (1957). Intravenous Infusion of Bone Marrow in Patients Receiving Radiation and Chemotherapy. N. Engl. J. Med. *257*, 491–496.

Tian, H., McKnight, S.L., and Russell, D.W. (1997). Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. Genes & Development *11*, 72–82.

Tiedt, R., Hao-Shen, H., Sobas, M.A., Looser, R., Dirnhofer, S., Schwaller, J., and Skoda, R.C. (2008). Ratio of mutant JAK2-V617F to wild-type Jak2 determines the MPD phenotypes in transgenic mice. Blood *111*, 3931–3940.

Tourmente, S., Chapel, S., Dreau, D., Drake, M.E., Bruhat, A., Couderc, J.L., and Dastugue, B. (1993). Enhancer and silencer elements within the first intron mediate the transcriptional regulation of the β3 tubulin gene by 20-hydroxyecdysone in Drosphila Kc cells. Insect Biochemistry and Molecular Biology *23*, 137–143.

Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C.N., Weiss, M.J., Crossley, M., and Orkin, S.H. (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. Cell *90*, 109–119.

Uchida, N. (1992). Searching for hematopoietic stem cells: evidence that Thy-1.1lo Lin- Sca-1+ cells are the only stem cells in C57BL/Ka-Thy-1.1 bone marrow. The Journal of Experimental Medicine *175*, 175–184.

Uhlen, M., Zhang, C., Lee, S., Sjostedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., et al. (2017). A pathology atlas of the human cancer transcriptome. Science *357*, 660–.

Ungureanu, D., Wu, J., Pekkala, T., Niranjan, Y., Young, C., Jensen, O.N., Xu, C.-F., Neubert, T.A., Skoda, R.C., Hubbard, S.R., et al. (2011). The pseudokinase domain of JAK2 is a dual-specificity protein kinase that negatively regulates cytokine signaling. Nature Structural & Molecular Biology *18*, 971–U21.

Usuki, F. (2000). Ataxia caused by mutations in the alpha -tocopherol transfer protein gene. Journal of Neurology, Neurosurgery & Psychiatry *69*, 254–256.

Vainchenker, W., BOUGUET, J., Guichard, J., and BretonGorius, J. (1979). Megakaryocyte Colony Formation From Human-Bone Marrow Precursors. Blood *54*, 940–945.

van den Oudenrijn, S., Bruin, M., Folman, C.C., Peters, M., Faulkner, L.B., de Haas, M., and Borne, von dem, A.E.G.K. (2000). Mutations in the thrombopoietin receptor, Mpl, in children with congenital amegakaryocytic thrombocytopenia. Br. J. Haematol. *110*, 441–448.

Van Esch, H., Groenen, P., Nesbit, M.A., Schuffenhauer, S., Lichtner, P., Vanderlinden, G., Harding, B., Beetz, R., Bilous, R.W., Holdaway, I., et al. (2000). GATA3 haplo-insufficiency causes human HDR syndrome. Nature *406*, 419–422.

Vigon, I., Mornon, J.P., Cocault, L., Mitjavila, M.T., Tambourin, P., Gisselbrecht, S., and Souyri, M. (1992). Molecular cloning and characterization of MPL, the human homolog of the v-mpl oncogene: identification of a member of the hematopoietic growth factor receptor superfamily. Pnas *89*, 5640–5644.

Vitrat, N., Cohen-Solal, K., Pique, C., Le Couedic, J.P., Norol, F., Larsen, A.K., Katz, A., Vainchenker, W., and Debili, N. (1998). Endomitosis of human megakaryocytes are due to abortive mitosis. Blood *91*, 3711–3723.

Wang, G.L., Jiang, B.H., Rue, E.A., and Semenza, G.L. (1995). Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. Pnas *92*, 5510–5514.

Wang, X., Crispino, J.D., Letting, D.L., Nakazawa, M., Poncz, M., and Blobel, G.A. (2002). Control of megakaryocyte-specific gene expression by GATA-1 and FOG-1: role of Ets transcription factors. The EMBO Journal *21*, 5225–5234.

Watt, A.J., Battle, M.A., Li, J.X., and Duncan, S.A. (2004). GATA4 is essential for formation of the proepicardium and regulates cardiogenesis. Pnas *101*, 12573–12578.

Wechsler, J., Greene, M., McDevitt, M.A., Anastasi, J., Karp, J.E., Le Beau, M.M., and Crispino, J.D. (2002). Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. Nat. Genet. *32*, 148–152.

Weidemann, A., and Johnson, R.S. (2009). Nonrenal regulation of EPO synthesis. Kidney Int *75*, 682–688.

Welte, T., Zhang, S.S.M., Wang, T., Zhang, Z., Hesslein, D.G.T., Yin, Z., Kano, A., Iwamoto, Y., Li, E., Craft, J.E., et al. (2003). STAT3 deletion during hematopoiesis causes Crohn's disease-like pathogenesis and lethality: a critical role of STAT3 in innate immunity. Pnas *100*, 1879–1884.

Wenger, R.H., and Hoogewijs, D. (2010). Regulated oxygen sensing by protein hydroxylation in renal erythropoietin-producing cells. AJP: Renal Physiology *298*, F1287–F1296.

Wiestner, A., Schlemper, R.J., van der Maas, A.P., and Skoda, R.C. (1998). An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia. Nat. Genet. *18*, 49–52.

Willekens, F.L.A., Werre, J.M., Groenen-Döpp, Y.A.M., Roerdinkholder-Stoelwinder, B., de Pauw, B., and Bosman, G.J.C.G.M. (2008). Erythrocyte vesiculation: a self-protective mechanism? Br. J. Haematol. *141*, 549–556.

Wilson, A., Laurenti, E., Oser, G., van der Wath, R.C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C.F., Eshkind, L., Bockamp, E., et al. (2008). Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. Cell *135*, 1118–1129.

Wilson, C.H., Gamper, I., Perfetto, A., Auw, J., Littlewood, T.D., and Evan, G.I. (2014). The kinetics of ER fusion protein activation in vivo. Oncogene *33*, 4877–4880.

Wirths, S., Bugl, S., and Kopp, H.-G. (2013). Steady-state neutrophil homeostasis is a demand-driven process. Cell Cycle *12*, 709–710.

Witthuhn, B.A., Quelle, F.W., Silvennoinen, O., Yi, T., Tang, B., Miura, O., and Ihle, J.N. (1993). JAK2 associates with the erythropoietin receptor and is tyrosine phosphorylated and activated following stimulation with erythropoietin. Cell *74*, 227–236.

Wolber, E.M., Fandrey, J., Frackowski, U., and Jelkmann, W. (2001). Hepatic thrombopoietin mRNA is increased in acute inflammation. Thromb. Haemost. *86*, 1421–1424.

Wu, H., Liu, X., Jaenisch, R., and Lodish, H.F. (1995). Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. Cell *83*, 59–67.

Xu, M., Li, J., Dias Neves, M.A., Zhu, G., Carrim, N., Yu, R., Gupta, S., Marshall, J., Rotstein, O., Peng, J., et al. (2018). GPIba is required for platelet-mediated hepatic thrombopoietin generation. Blood *132*, 622–634.

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., and Nakauchi, H. (2013). Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells. Cell *154*, 1112–1126.

Yamazaki, S., Souma, T., Hirano, I., Pan, X., Minegishi, N., Suzuki, N., and Yamamoto, M. (2013). A mouse model of adult-onset anaemia due to erythropoietin deficiency. Nat Commun *4*, 1950.

Yasuda, Y., Masuda, S., Chikuma, M., Inoue, K., Nagao, M., and Sasaki, R. (1998). Estrogen-dependent production of erythropoietin in uterus and its implication in uterine angiogenesis. J. Biol. Chem. *273*, 25381–25387.

Yoshihara, H., Arai, F., Hosokawa, K., Hagiwara, T., Takubo, K., Nakamura, Y., Gomei, Y., Iwasaki, H., Matsuoka, S., Miyamoto, K., et al. (2007). Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. Cell Stem Cell *1*, 685–697.

Yoshimura, A. (1998). The CIS family: Negative regulators of JAK-STAT signaling. Cytokine Growth Factor Rev. *9*, 197–204.

Youssoufian, H., Longmore, G., Neumann, D., Yoshimura, A., and Lodish, H.F. (1993). Structure, function, and activation of the erythropoietin receptor. Blood *81*, 2223–2236.

Yu, F., White, S.B., Zhao, Q., and Lee, F.S. (2001). HIF-1alpha binding to VHL is regulated by stimulus-sensitive proline hydroxylation. Pnas *98*, 9630–9635.

Zanjani, E.D., Ascensao, J.L., McGlave, P.B., Banisadre, M., and Ash, R.C. (1981). Studies on the liver to kidney switch of erythropoietin production. J. Clin. Invest. *67*, 1183–1188.

Zhang, B., Ng, D., Jones, C., Oh, S.T., Nolan, G.P., Salehi, S., Wong, W., Zehnder, J.L., and Gotlib, J. (2011). A novel splice donor mutation in the thrombopoietin gene leads to exon 2 skipping in a Filipino family with hereditary thrombocythemia. Blood *118*, 6988–6990.

Zmajkovic, J., Lundberg, P., Nienhold, R., Torgersen, M.L., Sundan, A., Waage, A., and Skoda, R.C. (2018). A Gain-of-Function Mutation in EPO in Familial Erythrocytosis. N. Engl. J. Med. *378*, 924–930.

JAKUB ZMAJKOVIČ

Belchenring 65, 4123 Allschwil, Switzerland
Phone: +41 77 96 100 10 E-mail: jakub.zmajkovic@gmail.com LinkedIn: http://www.linkedin.com/in/zmajkovic

Summary

- Strong interest and background in disease-related and translational research with the focus on cancer genetics, tumor biology, and hematology.
- Extensive hands-on experience with an array of biochemical, molecular, and cell-based techniques in an academic and industrial setting.
- Scientific excellence demonstrated by several publications in leading scientific journals, including first author publication in the NEJM.
- Problem-solving capabilities developer by planning, design, execution, and interpretation of scientific experiments on a daily basis.
- The capability of working in multidisciplinary teams as demonstrated by several fruitful international collaborations.

Professional experience

2014- present **Molecular mechanisms of myeloproliferative neoplasms**

Postdoctoral fellow/PhD candidate, Department of Biomedicine, University Hospital Basel

- Successfully established CRISPR/Cas9-mediated gene editing in the lab.
- Deciphered a novel molecular mechanism of hereditary erythrocytosis (with Anders Waage, Trondheim, Norway).
- Generated cell lines by CRISPR/Cas9 used for the development and screening of selective JAK2 inhibitors (with Olli Silvennoinen, Tampere, Finland).
- Supervised student assistant (6 months, funded by “Get on track” scholarship), summer student (3 months), and Praktikum student (2 months).

2012-2014 **Engineering and production of therapeutic monoclonal antibodies**

Research Assistant, mAb Engineering & Production, Arsanis Biosciences, Vienna

- Experimental design and execution of cloning strategies to create fully humanized IgG antibodies, followed by the expression in HEK/CHO cell systems.
- Research performed under GLP conditions and the company’s SOPs.
- Stringent protocol documentation into an online-patenting system.
- Strong collaboration between 3 pre-clinical development groups led to the selection of the lead molecule that is currently in Phase II of the clinical development.

2010-2012 **Cell polarity and asymmetric cell division in embryos and stem cells**

Research Fellow, Research Institute of Molecular Pathology (IMP), Vienna

- Examination of the role of proteins involved in cell polarity using C.elegans one-cell embryos using the state-of-the-art confocal spinning disc and wide-field light microscopy.

2007-2008 **HER2-positive breast cancer signaling and cell migration**

Internship /master thesis, Friedrich Miescher Institute for Biomedical Research (FMI), Basel

- Study of the mechanisms underlying breast cancer cell migration (master thesis).
- Established high-throughput imaging of transwell migration assays (with Markus Wartmann, NIBR Basel).

Education

2014-2018 Doctor of Philosophy, University of Basel, Switzerland

2008-2010 Master of Science with honors, Comenius University in Bratislava, Slovakia
- part of the studies at Faculty of Science, Katholieke Universiteit Leuven, Belgium (fellowship from the National Scholarship Agency of the Slovak Republic)

2004-2007 Bachelor of Science, Comenius University in Bratislava, Slovakia

Publications

Zmajkovic J, Lundberg P, Nienhold R, Torgersen ML, Sundan A, Waage A, Skoda RC. A Gain-of-Function Mutation in EPO in Familial Erythrocytosis. N Engl J Med. 2018 Mar 8;378(10):924-930.

Zmajkovic J, Pianta A, Nienhold R et al.: A mutation in the Kozak sequence of the THPO gene causes hereditary thrombocytosis. Manuscript in preparation.

Gu Z, Liu Y, Cai F, Cao H, **Zmajkovic J** et al.: Loss of EZH2 Activates BCAA Metabolism to Drive Myeloid Transformation. Cancer Discovery, accepted for publication.

Trstenjak N, Stulik L, Rouha H, **Zmajkovic J**, et al.: Adaptation of the Staphylococcus aureus leukocidin LukGH for the rabbit host by protein engineering. Biochem J. 2019 Jan 25;476(2):275-292.

Shimizu T, Kubovcakova L, Nienhold R, **Zmajkovic J** et al.: Loss of Ezh2 synergizes with JAK2-V617F in initiating myeloproliferative neoplasms and promoting myelofibrosis. J Exp Med. 2016 ;213(8):1479-96.

Szijártó V, Guachalla LM, **Zmajkovic J** et al.: Bactericidal Monoclonal Antibodies Specific to the Lipopolysaccharide O Antigen from Multidrug-Resistant Escherichia coli Clone ST131-O25b:H4 Elicit Protection in Mice. Antimicrob Agents Chemother. 2015 ;59(6):3109-16.

Rouha H, Badarau A, **Zmajkovic J** et al.: Five birds, one stone: neutralization of α-hemolysin and 4 bi-component leukocidins of Staphylococcus aureus with a single human monoclonal antibody. MAbs. 2015; 7(1):243-54.

Kucerova L*, **Zmajkovic J** *, Toro L, Skolekova S, Demkova L, Matuskova M.: Tumor-driven Molecular Changes in Human Mesenchymal Stromal Cells. Cancer Microenviron. 2014; 8(1):1-14.
* equal contribution

Awards and Achievements

2018 Department of Biomedicine Basel Research Prize for the best publication

2018 Bruno Speck Award (Stiftung für Hämatologische Forschung Basel)

2017 “Get on track” scholarship awarded by the University of Basel

2017 Abstract Achievement Award, ASH Annual Meeting 2017

2008 Slovak National Scholarship Agency fellowship to study at Katholieke Universiteit, Leuven

Courses and Skills

Laboratory skills:

- Course in Laboratory Animal Science FELASA Category B (2015)
- Techniques:
  - mammalian cell culture (adherent and non-adherent cell lines, primary cells), hematopoietic colony-forming cell assays
  - molecular cloning, CRISPR-mediated gene editing, transfections, single cell cloning
  - Western blotting, qRT-PCR, ELISA, multicolor flow cytometry,
  - mice handling, organ removal, serial blood sampling, i.p. and i.v. injections
- Databases for experimental animals (PyRAT) and biological samples (Spirit)
- Online documenting and patenting system (PatentSafe)

Soft skills:

- The Role of Intellectual Property in Life Sciences course (2015)
- Good Scientific Practice workshop (2017)
- Conflict management workshop (2018)
- Writing science articles for Slovak newspapers SME (authored 17 articles online and 12 in print)