Aus der
Berlin-Brandenburg School for Regenerative Therapies
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin
und der
Division of Hematology/Oncology, Boston Children's Hospital,
Harvard Medical School, Boston

DISSERTATION

Consequences of Ribosomal Perturbations in Human Hematopoiesis

zur Erlangung des akademischen Grades
Medical Doctor - Doctor of Philosophy (MD/PhD)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von
Rajiv Kumar Khajuria
aus Berlin

Datum der Promotion: 04.06.2021

Table of Contents

1. Abstract

Hematopoiesis, the formation of blood cells, is probably the best understood system of cellular differentiation in mammalian biology. Traditionally, this process is thought to occur as a hierarchical progression of differentiation from hematopoietic stem cells (HSCs), where progressively lineage restricted oligo-, bi-, and unipotent progenitors are generated, eventually producing mature blood cells. However, this classical view is based on analyses performed in marker pre-defined bulk cell populations and has been challenged. Recent studies, enabled by single cell technologies, suggest that lineage commitment in hematopoiesis occurs much earlier than previously thought, at the hematopoietic stem and progenitor cell (HSPC) stage. It remains unclear what relevance these refined models of hematopoiesis have with regard to human diseases. Furthermore, while tremendous insight has been gained into this process on the transcriptional level, the contribution of post-transcriptional regulation in hematopoiesis remains to be largely explored. Here, we use an experiment of nature, the rare congenital disorder Diamond-Blackfan anemia (DBA), where the majority of mutations affect ribosomal proteins and the erythroid lineage is selectively perturbed, as a model to address these issues.

We use human genetics to gain deeper insight into the ribosomal defects in DBA in vivo, biochemical and proteomic studies to examine ribosome levels and composition in human hematopoietic cells with DBA-associated molecular lesions, ribosome profiling in HSPCs undergoing erythroid lineage commitment to assess changes in translation globally, transcriptome analyses of hematopoietic master regulators from unperturbed human HSPCs, and intracellular flow cytometry analyses of primary DBA patient samples to elucidate the pathological mechanisms underlying DBA and how these relate to physiologic lineage commitment.

We find that in DBA, the cellular levels of ribosomes are reduced, while the ribosome protein composition remains unaltered. This global reduction of ribosome levels impairs the translation of a select subset of transcripts including the key erythroid transcription factor GATA1. These transcripts have, among other properties, shorter and less structured 5' UTRs than unaffected mRNAs, features that *GATA1* exhibits relatively uniquely among hematopoietic master regulators, which may explain the erythroid

lineage selective defect in DBA. Finally, we show that GATA1 protein levels are reduced already at the HSPC stage in primary DBA patients' bone marrow specimens, which fits to the refined models of hematopoiesis that suggest lineage commitment takes place in this primitive cell compartment.

By studying the rare congenital disorder DBA, we gain insight into how cellular ribosome levels and translation play a key role in the process of human hematopoietic lineage commitment.

1.1 Abstract (German)

Die Hämatopoese, die Bildung von Blutzellen, ist wahrscheinlich das am besten verstandene System der zellulären Differenzierung in der Biologie von Säugern. Traditionell nimmt man an, dass dieser Prozess als ein von hämatopoetischen Stammzellen (HSZ) ausgehendes, hierarchisches Fortschreiten der Zelllinendifferenzierung geschieht, wobei zunehmend Zelllinien-beschränkte oligo-, bi-, und unipotente Progenitorzellen generiert werden, die schließlich reife Blutzellen produzieren. Dieses klassische Konzept basiert jedoch auf Analysen, die in Marker prädefinierten Mischzellpopulationen durchgeführt wurden und wird nun in Frage gestellt. Neuere Studien, ermöglicht durch Einzelzelltechnologien, suggerieren, dass die Zelllinienspezifizierung früher als bisher gedacht, nämlich in hämatopoetischen Stamm- und Progenitorzellen (HSPZ) stattfindet. Welche Relevanz diese weiterentwickelten Modelle der Hämatopoese mit Hinblick auf humane Erkrankungen haben, ist unklar. Außerdem, während enorme Einblicke in diesen Prozess auf dem Transkriptionslevel gewonnen werden konnten, ist die Rolle der post-transkriptionellen Regulation in der Hämatopoese größtenteils unerforscht. Hier nutzen wir ein „Experiment der Natur“, die seltene kongenitale Erkrankung Diamond-Blackfan Anämie (DBA), welche in der Mehrheit der Fälle durch Mutationen in ribosomalen Proteinen verursacht wird und wobei die erythroide Zelllinienentwicklung selektiv gestört ist, als ein Modell um diese Aspekte zu adressieren.

Wir nutzen humangenetische Untersuchungen um tiefere Einblicke in die ribosomalen Defekte in DBA in vivo zu erlangen, biochemische und proteomische Studien um die Ribosomenlevel und Ribosomenproteinkomposition in humanen hämatopoetischen Zellen mit DBA-assoziierten Molekularläsionen zu ermitteln, Ribosome Profiling in sich erythroider Differenzierung unterziehenden HSPZ um globale Änderungen in der Translation zu evaluieren, Transkriptomanalysen von hämatopoetischen Schlüsselregulatoren in nicht-pathologischen HSPZ und intrazelluläre Durchflusszytometrieanalysen von primären DBA Patientenproben um die der DBA zugrunde liegenden Pathomechanismen zu ergründen und diese in Bezug zur physiologischen Zelldifferenzierung zu stellen.

Wir demonstrieren, dass die zellulären Ribosomenlevel in DBA reduziert sind, die Ribosomenproteinkomposition jedoch unverändert ist. Diese globale Reduktion der Ribosomenlevel beeinträchtigt die Translation einer bestimmten Untergruppe von Transkripten einschließlich des erythroiden Schlüssel-Transkriptionsfaktors GATA1. Diese Transkripte haben unter anderem kürzere und weniger strukturierte 5' UTRs als unbeeinträchtigte mRNAs, Eigenschaften, die *GATA1* vergleichsweise einzigartig unter hämatopoetischen Schlüsselregulatoren aufweist, was den die spezifisch erythroide Zellliniendifferenzierung betreffenden Defekt in DBA erklären könnte. Abschließend zeigen wir, dass die GATA1 Proteinlevel schon auf der Stufe der HSPZ in primären DBA-Patientenknochenmarksproben reduziert sind, was zu den weiterentwickelten Modellen der Hämatopoese passt, die suggerieren, dass die Zelllinienspezifizierung schon in diesem primitiven Zellkompartment stattfindet.

Durch die Untersuchung der seltenen kongenitalen Erkrankung DBA erhalten wir Einblick darin, wie zelluläre Ribosomenlevel und Translation eine Schlüsselrolle in dem Prozess der humanen hämatopoetischen Zellliniendifferenzierung spielen.

2. Introduction

This thesis is based on the original publication with the title "Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopiesis"[1]. Accordingly, formulations throughout the text may be partly or completely adopted from the original manuscript.

2.1 Lineage Commitment in Hematopoiesis

Blood cell production or hematopoiesis is one of the best-understood examples of cellular differentiation in mammalian physiology[2]. Most of our knowledge of this process stems from analyses done in marker pre-defined bulk cell populations. This previous work suggested a hierarchical progression of differentiation proceeding from hematopoietic stem cells (HSCs), where increasingly more lineage-restricted oligo-, bi-, and unipotent progenitors are produced, eventually forming mature circulating blood cells[2,3]. This traditional view of hematopoiesis has recently been challenged by studies enabled through single-cell technologies, which suggest that lineage commitment occurs much earlier, in hematopoietic stem and progenitor cells (HSPCs), which then undergo orderly differentiation into mature circulating blood cells[4-6]. These studies revealed that HSPCs are transcriptionally and epigenetically more heterogeneous than previously recognized and concurrently exhibit biases in cell fate decisions. The vast majority of our molecular knowledge of these differentiation processes relies on work performed on the transcriptional level[4,5], while it is clear that post-transcriptional regulation is also critical for both stem cell maintenance and cell differentiation. The significance of such post-transcriptional regulation is emphasized by the fact that cellular protein abundances are only partly explained by cellular mRNA levels for a given protein[7]. While considerable insight has been gained into the role of post-transcriptional regulation in the maintenance of hematopoietic stem cells (HSCs)[8], and suggests significant changes in protein synthesis rates in early stages of lineage commitment, the post-transcriptional regulation of molecular regulators of these differentiation processes and functional consequences of

perturbations of these remain largely unexplored, also in processes beyond hematopoiesis.

2.2 Ribosomopathies Including Diamond-Blackfan Anemia

There is an increasing appreciation for the key roles of post-transcriptional regulation in cellular differentiation processes, such as hematopoiesis[8]. Strikingly, naturally-occurring mutations affecting protein synthesis cause distinct and cell-type specific defects[9]. Mutations in ubiquitously expressed ribosomal proteins cause highly-penetrant human diseases, the so-called ribosomopathies, with exquisitely specific phenotypes: Diamond-Blackfan anemia (DBA)[10], congenital asplenia[11], and neurodevelopmental disorders[12]. It remains a mystery how such specific defects emerge from mutations in proteins that are required by every cell[9]. In the case of DBA, a rare congenital blood disorder, there is a selective impairment in the production of erythroid precursors and progenitors in the bone marrow of patients, while all other hematopoietic lineages appear unaffected in their development[13,14]. The majority of DBA cases are caused by heterozygous loss-of-function mutations in one of at least 26 different ribosomal protein (RP) genes, resulting in RP haploinsufficiency[15]. Why mutations in proteins so ubiquitous as RPs selectively affect the erythroid lineage while sparing all other hematopoietic lineages, has remained incompletely understood. We recently identified the first non-RP mutations in rare DBA patients in the hematopoietic master transcription factor GATA1[16], which is critical for erythroid differentiation. Subsequently, we were able to link these two seemingly disparate pathologies together by discovering that RP haploinsufficiency results in selectively impaired translation of *GATA1* mRNA, as compared to a number of other important erythroid factors[10]. How this selective impairment of translation arises from RP haploinsufficiency, remains to be explored.

2.3 Ribosome Defect Models

A number of models for how exactly RP haploinsufficiency causes a selective defect in translation and tissue-specific phenotypes have been proposed[17] including the two most prominent, but not necessarily mutually exclusive, models: The “qualitative” ribosomal

defect model posits that under physiologic conditions, the ribosome pool is heterogeneous in terms of RP composition and the resulting "specialized" ribosomes that are qualitatively distinct preferentially translate specific mRNAs. Therefore, RP haploinsufficiency would promote the production of ribosomes lacking one or more specific RPs and would thereby alter protein translation. This concept is supported by recent studies in a non-DBA related Rpl38 haploinsufficient mouse model with homeotic transformation and skeletal patterning defects due to altered translation of a select subset of *Hox* mRNAs[18]. Conversely, the "quantitative" ribosomal defect model posits that reduced levels of ribosomes of *equivalent* protein composition would cause the mRNA pool to compete for the limited amount of ribosomes and thereby lead to decreased ribosome association for a subset of mRNAs[19].

2.4 Aims of the Project and Thesis

We consider DBA as an experiment of nature, studying this disorder gives us the opportunity to understand how perturbations in the ribosome can result in a selective impairment of translation thereby selectively impairing erythroid lineage commitment, while allowing other hematopoietic lineages to be normally produced. Thus, studies of DBA do not only allow us to gain insight into the pathomechanism of this disease, but also provide us a better understanding of the role of post-transcriptional regulation in lineage commitment in human hematopoiesis more generally, potentially with implications also in other tissues. The major aims of this project were to

1. Identify previously unknown mutations causing DBA to gain more insight into the pathologic mechanisms of the disease.
2. Assess ribosome levels and ribosome protein composition in human cellular models of DBA in order to discriminate between the aforementioned qualitative and quantitative ribosome defect models that underlie the altered translation and lineage specific defect in this disorder.
3. Identify the transcripts whose translation is down-regulated by ribosomal defects seen in DBA on a genome-wide scale.

4. Characterize these translationally affected transcripts including their 5' untranslated regions (UTRs) in order to better understand why only a subset of mRNAs is impaired in its translation by the ribosome defects that cause DBA.
5. Use bone marrow specimens from DBA patients in order to identify the stages within hematopoiesis at which the translational defects occur and thus being able to assess the relevance of the above mentioned refined models of hematopoiesis with regard to human disease.

3. Methods

A comprehensive, detailed and complete description of the methods and materials as well as of the bioinformatics analyses can be found in the STAR METHODS and tables section of the original publication[1]. The key resources table in the original publication contains also all relevant catalogue numbers. Due to space constraints, here, I mostly focus on the description of the methodologies that could not be presented in detail in the original publication.

Ribosome Profiling

Harvesting of Samples

Cells were incubated with 100 mg/ml of cycloheximide (Sigma Aldrich) for 5 minutes at 37°C, washed twice with ice-cold PBS containing 100 mg/ml of cycloheximide and lysed in 10 mM Tris-HCl (pH 7.4), 5 mM MgCl2, 100 mM KCl, 1% Triton X-100, 3 mM DTT, 100 mg/ml cycloheximide, 500 U/ml RNasin (Promega) and 1 x Complete Protease Inhibitor, EDTA-free (Roche) as well as 1x Protease Inhibitor Set (without EDTA) (G-Biosciences). Lysates were partitioned for either ribosome footprint profiling or mRNA sequencing.

Total RNA Extraction

Total RNA was extracted with the Direct-zol™ RNA MiniPrep Plus w/ TRI Reagent® Kit (Zymo Research). To this end, 3 volumes of TRI Reagent® were added to each lysate partitioned in the step described above for mRNA isolation in the further process and mixed thoroughly. For RNA purification, an equal volume of ethanol (95 – 100 %) was added to each sample lysed in TRI Reagent® and mixed thoroughly. The mixture was transferred into a Zymo-Spin™ IIICG column in a collection tube, respectively and centrifuged. Reloading of the spin column was required to process the complete volume. The spin column was transferred into a new collection tube and the flow-through was discarded, respectively. For DNaseI treatment, 400 µl of RNA Wash Buffer were added to each column and then centrifuged. For each column, 5 µl DNase I (6 U/µl) were added to 75 µl of DNA Digestion Buffer in a RNase-free tube and mixed. This mix was added

directly to the column matrix, respectively and incubated at room temperature for 15 minutes. Subsequently, 400 µl Direct-zol™ RNA PreWash were added to each column and centrifuged. The flow-through was discarded, respectively and the step was repeated. Next, 700 µl RNA Wash Buffer were added to each column and centrifuged for 2 minutes in order to completely remove the wash buffer. The columns were then transferred to a new RNase-free tube, respectively and total RNA was eluted by adding 100 µl of nuclease-free water directly onto the column matrix and centrifugation. If not used immediately for further processing, total RNA was flash frozen and stored at -80°C.

mRNA Isolation by Poly(A) Selection

Total mRNA was poly-A selected using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (New England Biolabs). Accordingly, 50 µl of the eluted total RNA/sample from the step described before were added to a nuclease-free 200 µl PCR tube. In a separate PCR tube, 20 µl of well resuspended NEBNext Magnetic Oligo d(T)25 Beads were added. The beads were washed by adding 100 µl of RNA Binding Buffer and mixed thoroughly by pipetting the entire volume up and down at least 6 times. The tubes with the beads were then placed on a magnetic rack at room temperature until the solution was clear, but at least for 2 minutes. The supernatant was then completely removed from the tubes and discarded without disturbing the beads. This washing step was repeated at least once. The beads were then resuspended in 50 µl of RNA Binding Buffer and the 50 µl of total RNA/sample were added to a tube with resuspended beads, respectively. The entire volume was thoroughly mixed by pipetting up and down. Next, the samples were heated on a thermal cycler at 65°C for 5 minutes and held at 4°C in order to denature the RNA and facilitate binding of the poly-A-RNA to the beads. The tubes were then removed from the thermal cycler and the beads thoroughly mixed by pipetting up and down for at least 6 times. The samples were then placed on the bench for 5 minutes at room temperature in order to allow the RNA to bind to the beads. This step with thorough mixing of the beads followed by the 5 minutes of incubation at room temperature was repeated one time. The tubes were then placed on a magnetic rack for at least 2 minutes in order to separate the poly-A RNA bound to the beads from the solution. The supernatant was removed and discarded without disturbing the beads. The

tubes were then removed from the magnetic rack and the beads were washed with 200 µl of Wash Buffer to remove unbound RNA, respectively. The beads were mixed thoroughly by pipetting up and down the entire volume at least 6 times and the tubes were placed on a magnetic rack again for at least 2 minutes. All supernatant was removed and discarded, respectively without disturbing the beads. This washing step was repeated one more time. Then, 50 µl of Tris Buffer were added to each tube and samples were mixed thoroughly by pipetting up and down at least 6 times. The tubes were then placed on a thermal cycler and heated at 80°C for 2 minutes and then held at 25°C to elute the poly-A RNA from the beads. The tubes were then removed from the thermal cycler and 50 µl of RNA Binding Buffer were added to each tube to allow the RNA to bind to the same beads. The samples were mixed thoroughly by pipetting up and down at least 6 times. The tubes were then incubated at room temperature for 5 minutes in order to allow the RNA to bind to the beads. This last step with thorough mixing and 5 minutes of incubation was repeated one more time. Next, the tubes were placed on the magnetic rack again for at least 2 minutes. Then, all supernatant was removed and discarded without disturbing the beads, respectively. The tubes were then removed from the magnetic rack again and 200 µl of Wash Buffer were added to each sample, which was then thoroughly mixed by pipetting up and down 6 times. The tubes were then once again placed on a magnetic rack for at least 2 minutes and the supernatant was removed and discarded, respectively without disturbing the beads. The tubes were then removed from the magnetic rack and mRNA was eluted from the beads by adding 20 µl of Tris Buffer, thorough mixing by pipetting up and down at least 6 times and incubation of the samples at 80°C for 2 minutes, then held at 25°C. The tubes were then immediately placed on a magnetic rack for at least 2 minutes and the supernatant containing the eluted mRNA was transferred to new nuclease-free tubes. If not used immediately for further processing, mRNA was flash frozen and stored at -80°C. mRNA seq libraries were generated as described previously[20].

Ribosome Footprinting

Ribosome footprinting and subsequent library preparation of ribosome protected RNA fragments (RPFs) was performed with the Truseq Ribo Profile (Mammalian) Kit

(Illumina) with some modifications. Accordingly, in 200 µl of the partitioned lysate described above under "Harvesting of Samples", RNase I (Ambion) digestion was done at a concentration of 2.5 U/µl lysate for 45 minutes at room temperature with gentle mixing. During this incubation, MicroSpin S-400 columns (GE Healthcare Life Sciences) were prepared for RPF purification. 3 ml of 1X Mammalian Polysome Buffer/sample were prepared by combining 600 µl of 5X Buffer with 2.4 ml of nuclease-free water. The MicroSpin S-400 columns were inverted several times to resuspend the resin. Each column was opened on both ends to allow the buffer to drip out under gravity. Next, a collection tube was attached to each column and centrifuged for 4 minutes at 600 × g in a fixed-angle, benchtop centrifuge at room temperature. The flow-through was discarded and each column was transferred to a 1.5 ml tube. The RNase I digestion reactions were stopped by adding 15 µl of SUPERase•In RNase Inhibitor and chilling the samples on ice, respectively. Before allowing the column to dry, 100 µl of nuclease-digested RPF sample were applied, respectively and centrifuged for 2 minutes at 600 × g. The flow-through was collected and 10 µl of 10% SDS were added to each sample. For purification of RPF RNA, 220 µl of RNA Binding Buffer and 495 µl of 100 % ethanol were added to each sample and mixed thoroughly. The further procedure was performed according to the RNA Clean & Concentrator-25 Kit method (Zymo Research). Samples were eluted in 25 µl of nuclease-free water, respectively.

rRNA Depletion

rRNA removal was performed by using the Ribo-Zero Gold rRNA Removal Kit (Illumina). For each Ribo zero reaction, 225 µl of magnetic beads were required at room temperature and thoroughly vortexed for homogeneity. The beads containing tube was then placed on a magnetic rack for at least 2 minutes for the batch washing procedure. The supernatant was removed and discarded without disturbing the beads. The tube was then removed from the magnetic rack and an equal amount of RNase-free water was added. The tube was then thoroughly vortexed. This bead washing step was repeated one more time. The tube was then placed on the magnetic rack again for at least 2 minutes and the supernatant was removed and discarded. The tube was removed from the magnetic rack and a volume of Magnetic Bead Resuspension equal to the number of

reactions x 60 µl was added to the tube, that was then thoroughly vortexed. 65 µl of the washed and resuspended magnetic beads per reaction were then aliquoted into 1.5 ml RNase-free microcentrifuge tubes. To each tube of resuspended beads, 1 µl of RiboGuard RNase Inhibitor was added, mixed briefly and vortexed. The washed beads were then stored at room temperature until needed in further steps. In a RNase-free tube, per sample 16 µl of RNase-free water, 4 µl of Ribo-Zero rRNA Reaction Buffer, 10 µl of RNA sample and 10 µl of Ribo-Zero Removal Solution were mixed thoroughly by pipetting the entire volume 10-15 times. This mix was then incubated at 68°C for 10 minutes, removed from the heat and then centrifuged briefly to collect any condensation. The tubes were then incubated for 5 minutes at room temperature. 40 µl of the now probe-hybridized sample were added to the 65 µl of washed, room temperature magnetic beads and mixed immediately thoroughly by pipetting the entire volume up and down for 10-15 times. This mixture step was repeated one more time. The tubes were then capped and vortexed at high speed for at least 10 seconds, then incubated at room temperature for 5 minutes. Subsequently, the tubes were placed on a magnetic rack for at least 2 minutes and the supernatant containing the depleted rRNA sample was transferred to RNase-free microcentrifuge tube. The Ribo-Zero-treated RNA was then purified by using a modified RNA Clean & Concentrator-5 Kit (Zymo Research) method. Each sample was brought up to a volume of 100 µl. 200 µl of RNA Binding Buffer and 450 µl of 100 % ethanol were added to each sample and mixed thoroughly. The remaining procedure was then performed according to the manufacturer's instructions, the purified RPF RNA was eluted in 11 µl of nuclease-free water.

PAGE Purification of RPFs

5 µl of TruSeq Ribo Profile RNA Control (28 nt = 5' NNGUACACGGAGUCGACCCGCAACGCNN 3'; 30 nt = 5' NNGUACACGGAGUCAAGACCCGCAACGCNN 3') were added to a 0.5 ml microcentrifuge tube together with 5 µl of denaturing gel loading dye, respectively. For each RPF RNA sample, 10 µl of the RNA were mixed with 10 µl of denaturing gel loading dye. A mix of 4 µl of the 20/100 ladder, 1 µl of water and 5 µl of denaturing loading dye was prepared to prevent cross contamination for loading between RPF RNA

samples. The ladder and samples were denatured at 95°C for 5 minutes and then immediately placed on ice. 10 µl of ladder and each sample were loaded onto a 12% urea-polyacrylamide gel and run at 180 V until the Bromophenol Blue band reached the bottom of the gel (about 210 V for 1 hour). The gel was then stained with SYBR Gold at 4°C and RNA was visualized with a dark-field transilluminator. The gel slices corresponding to the 28 nt and 30 nt of the TruSeq Ribo Profile RNA Control were excised for each sample and transferred to a 0.5 ml microcentrifuge tube, that had a hole punched in the bottom with a sterile 20 gauge needle and was placed in a 1.5 ml centrifuge tube. The samples were then centrifuged for 2 minutes at 12,000 x g to shred the gel slices and the 0.5 ml tubes were discarded. To each sample, 400 µl of nuclease-free water, 40 µl of 5 M Ammonium Acetate and 2 µl of 10 % SDS were added and the tubes then gently shook at room temperature for 2 hours to elute the RNA from the gel. Then, a 1 ml pipette with a trimmed tip was used to transfer the slurry to 1.5 ml filter tubes, which were then centrifuged for 3 minutes at 2,000 x g in order to separate the disrupted gel slices from the eluted RNA. Next, the aqueous solution was transferred into new 1.5 ml tubes, respectively and 2 µl of Glycogen and 700 µl of 100 % isopropanol were added to each tube, which were then stored at -20°C for at least 1 hour. Subsequently, the tubes were centrifuged at 4°C for 20 minutes at 12,000 x g in order to pellet the RPF RNA. The pellet was then washed with 80 % ethanol and air dried, respectively. Finally, each RPF RNA sample pellet was resuspended in 20 µl of nuclease-free water. If not immediately used for further procedures, the RNA was flash frozen and kept at -20°C.

End Repair

In order to prepare the RPF RNA for 3'adapter ligation, samples were end-repaired. To this end, 7.5 µl of TruSeq Ribo Profile PNK Buffer were added to 20 µl of each sample kept on ice. Next, to each RPF RNA sample 44.5 µl of nuclease-free water and 3 µl of TruSeq Ribo Profile PNK were added, mixed well and incubated at 37°C for 1 hour. Each sample volume was then added up to 100 µl with 25 µl of nuclease-free water and purified with a modified RNA Clean & Concentrator-5 Kit (Zymo Research) method as described above. Finally, samples were eluted in 11 µl of nuclease-free water.

3'Adapter Ligation

To each sample kept on ice, 1 µl of TruSeq Ribo Profile 3' Adapter (5' AGATCGGAAGAGCACACGTCT 3') was added and samples were then heat-denatured in a thermal cycler for 2 minutes at 65°C, then held at 4°C. PAGE purified TruSeq Ribo Profile RNA Control was used as positive control. In the meantime, a ligation mastermix was prepared with 3.5 µl of TruSeq Ribo Profile Ligation Buffer, 1 µl of 100 mM DTT and 1.5 µl of TruSeq Rubo Profile Ligase per sample. 6 µl of this ligation mix were added to each denatured RNA sample, mixed thoroughly by pipetting the entire volume up and down multiple times, centrifuged briefly and incubated at room temperature for 2 hours. Then, 2 µl of TruSeq Ribo Profile AR Enzyme were added to each reaction, mixed thoroughly and incubated at 30°C for 2 hours.

Reverse Transcription, cDNA PAGE Purification and cDNA Circularization

A reverse transcription mastermix was prepared with 4.5 µl of TruSeq Ribo Profile RT Reaction Mix, 1.5 µl of 100 mM DTT, 6 µl of nuclease-free water and 1 µl of EpiScript RT for each reaction. 13 µl of this mix were added to each reaction, mixed well and incubated for 30 minutes at 50°C in a thermal cycler with a heated lid. Then, 1 µl of TruSeq Ribo Profile Exonuclease was added to each reaction, which were then incubated at 37°C for 30 minutes, then 80°C for 15 minutes and then held at 4°C. Subsequently, 1 µl of TruSeq Ribo Profile RNase Mix was added to each reaction, mixed well, incubated at 55°C for 5 minutes and then held at 4°C. 18 µl of nuclease-free water were added to each reaction to adjust the volume to 50 µl, respectively. Samples were then purified by using a modified RNA Clean & Concentrator-5 Kit (Zymo Research) method and eluted in 11 µl of nuclease-free water. PAGE Purification of the cDNA was then performed analogous to the RPF PAGE purification described above with some slight modifications and the cDNA was eluted in 11 µl of nuclease-free water. For circularization of the cDNA, a mastermix for each sample was prepared consisting of 4 µl of TruSeq Ribo Profile CL Reaction Mix, 2 µl of ATP, 2 µl of $MnCl_2$ and 2 µl of CircLigase. 10 µl of this mix were added to each reaction, which were then mixed gently and centrifuged briefly. Next, these reactions were incubated at 60°C for 2 hours, then held at 4°C.

PCR Amplification and PCR Product PAGE Purification

A PCR mastermix for each sample was prepared with 16 µl of nuclease-free water, 2 µl of TruSeq Ribo Profile Forward PCR Primer (5' AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACG 3'), 2 µl of TruSeq Ribo Profile Index PCR Primer and 25 µl of 2X Phusion Master Mix. 45 µl of this mix were added to each 5 µl of undiluted cDNA sample, mixed well and run on a thermal cycler at 98°C for 30 seconds, 10 cycles at 94°C for 15 seconds, at 55°C for 5 seconds, at 65°C for 10 seconds, then held at 4°C. The PCR products were then purified using 90 µl (1.8X) of Agencourt AMPure XP beads according to the manufacturer's instructions. The libraries were eluted in 16 µl of nuclease-free water. The PCR products were loaded with native gel loading dye containing Bromophenol Blue onto 8% native polyacrylamide gels in 1X TBE for PAGE purification. Gel lanes containing samples were separated by lanes containing 5 µl of 20 bp ladder. The gel was run then at 200 V for ~30 minutes until the Bromophenol Blue reached the bottom of the gel, which was then stained with SYBR Gold. Libraries were visualized under a dark field transilluminator with an expected size of 140-160 bp, respectively. The further procedure was performed analogous to the RPF PAGE purification described above with some modifications. Libraries were resuspended in nuclease-free water and checked on a Bioanalyzer using the Agilent High Sensitivity DNA Assay. All libraries were sequenced on a HiSeq 2500 system (Illumina).

p53 Western Blotting and Gene Set Enrichment Analysis

The general western blotting procedure performed was analogous to those described in detail for detection of other proteins in the original publication[1]. In this process, membranes were blocked with p53 rabbit polyclonal antibody (FL-393, sc-6243, Santa Cruz Biotechnology) at a 1:1,000 dilution and incubated with donkey anti-rabbit peroxidase-coupled secondary antibody (711-035-152, Jackson ImmunoResearch) at a 1:10,000 to 1:20,000 dilution.

Gene set enrichment analysis (GSEA) was used with the "Preranked" option and 10,000 permutations for ΔmRNA[21]. The p53 gene set was obtained from the "Hallmark" gene sets provided by MSigDB (http://software.broadinstitute.org/gsea/msigdb).

Heme Quantification

A total of 200,000 cells per replicate from Luc, TSR2, RPS19 or RPL5 shRNA treated $CD34^+$ cells differentiated towards the erythroid lineage were harvested on day 5 following infection for heme quantification using the QuantiChrom Heme Assay Kit (BioAssay Systems; DIHM-250). The assay was performed in triplicates, following the manufacturer's protocol.

Statistical Analyses

All pairwise comparisons were assessed using an unpaired two-tailed Student's *t*-test, unless otherwise indicated in the main text or in the figure legends. Results were considered significant if the *P* value was <0.05.

Accession Codes

The raw mass spectrometry data have been deposited in the public proteomics repository MassIVE: MSV000080283.

The RNA-seq and ribosome profiling data reported in this thesis and in the original manuscript are deposited in the Gene Expression Omnibus (GEO) data repository: GSE89183.

4. Results

References to figures in this result section correspond to the figures in the original publication[1]. Due to space limitations, I focused on the key results and did not refer to every single figure panel of the publication.

4.1 DBA Mutations in *TSR2* Highlight the Importance of Ribosome Production in Hematopoiesis

We reasoned that the further identification of novel genetic causes of DBA would provide additional insight into pathologic mechanisms of this disorder. By performing whole exome sequencing, we and others[22] identified a novel hemizygous missense mutation in the X-linked *TSR2* gene in two male cousins with DBA, who lacked pathogenic mutations in all known DBA genes (**Figures 1A** and **S1A**). Similar to the known RP genes and *GATA1*, *TSR2* showed very few nonsense or frameshift mutations in 60,706 control individuals (**Figure S1A**). This was an important finding considering that TSR2 has been shown to act as chaperone for the yeast ortholog of RPS26 (eS26 in revised RP nomenclature) and to be crucial for the formation of the mature ribosome, while being completely localized in the nucleus[23], indicating that it is unlikely to interact with the mature ribosome. We could verify that the identified mutation is indeed a loss-of-function allele (**Figure 1B**) and also that TSR2 is completely localized to the nucleus in human hematopoietic cells (**Figure S1B**).

Next, we wanted to assess the consequences of TSR2 perturbations in human erythropoiesis. TSR2 suppression by short hairpin RNAs (shRNAs) (**Figure 1C**) resulted in an impaired erythroid differentiation from primary human HSPCs (**Figures 1D** and **S1C**) and an increased apoptosis phenotype, impaired growth, and a less mature erythroid gene signature (**Figures S1D-F**) consistent with phenotypes we observed previously in our *in vitro* primary cell models of DBA due to RP haploinsufficiency[10]. Also consistent with our previous DBA studies due to RP haploinsufficiency, TSR2 suppression resulted in selectively reduced GATA1 protein levels, as compared to other erythroid-important factors, including EPOR, STAT5A, and JAK2, while *GATA1* mRNA levels appeared to

be unaffected (**Figures 1E**, **1F** and **S1H-S1J**). Importantly, increased GATA1 protein expression could rescue erythroid lineage commitment from HSPCs with TSR2 suppression (**Figures 1G**, **1H**, **S1K** and **S1L**).

4.2 Molecular Lesions Underlying DBA Reduce Ribosome Levels in Hematopoietic Cells

Since the yeast ortholog of TSR2 has been shown to function as a key ribosome biogenesis factor[23], we interrogated whether human TSR2 has a similar function. Analogous to the yeast ortholog and also to the role of RPs in the maturation of ribosomes, suppression of human TSR2 results in impaired processing of rRNA (**Figures 2A**, **S2A** and **S2B**), strongly suggesting that human TSR2 is indeed critical for the levels of mature ribosomes in the cytoplasm that are available for translation. Indeed, these defects are often observed in DBA patients, where they can aid in diagnosis[24]. Next, we sought to examine the levels of actively translating ribosomes in both primary hematopoietic cells and cell lines with DBA-associated molecular lesions including suppression of TSR2, RPS19 (eS19), RPL5 (uL18), RPS24 (eS24), and RPL11 (uL5). Using quantitative polysome profiling, we consistently observed a reduced abundance of assembled monosomes and actively translating polysomes (**Figures 2D-2F** and **S2C-S2G**). Consistent with these findings, not only did we observe that the total protein abundance of the targeted RP was reduced, but a number of non-targeted RPs from both subunits, particularly among those found in the same subunit as the primary molecular lesion, were reduced in hematopoietic cells with DBA-associated lesions (**Figures 2B**, **2C** and **S2H-S2Q**). The decreased levels of cellular ribosomes in hematopoietic cells with DBA-associated lesions, together with the fact that pathogenic mutations in at least 26 distinct RPs[15] and *TSR2*[22] have been implicated in DBA, provides some initial support for the quantitative ribosome defect model underlying the altered translation in this disorder and for a selective role of ribosome levels in hematopoietic lineage commitment. However, these data cannot exclude a potential additional qualitative defect in the ribosome protein composition.

4.3 Verification of Constant Ribosome Composition in Human Hematopoietic Cells with DBA-Associated Molecular Lesions

Recent work has suggested that lesions in RPs may result in altered translation due to altered ribosome protein composition in some contexts[18]. To directly interrogate the protein composition of actively translating ribosomes in hematopoietic cells with DBA-associated lesions, we performed quantitative high-coverage tandem-mass-tag (TMT) mass spectrometry to measure the expression of all RPs. We fractionated cells by sucrose gradient sedimentation and collected monosomes (M, a single ribosome), light polysomes (LPs, 2-4 ribosomes), and heavy polysomes (HPs, ≥ 5 ribosomes) (**Figure 3A**). The RP composition within monosomes, light polysomes, and heavy polysomes was largely constant between controls and cells with DBA-associated lesions (**Figures 3B-3D**). The protein expression of the targeted RP did not deviate significantly from that of the non-targeted RPs (based on studentized residuals, **Figures S3G–S3I**). These data strongly suggest that the translational defects and the perturbation in erythroid lineage commitment in DBA are consequences from reduced ribosome levels, rather than from ribosomes with an altered protein composition. We note that, because our assay measures total protein levels within a given cellular fraction, we cannot exclude the possibility that the pool of actively translating ribosomes is comprised of ribosomes with multiple compositions and the relative abundances of these variable ribosomes is altered in cells with DBA-associated molecular lesions. However, it seems unlikely that this possibility would result in the same *total* RP composition in normal and perturbed cells (**please see also under 5.2 in the Discussion and Figure S9**).

4.4 Defining Transcripts Whose Translation Is Most Sensitive to DBA-Associated Molecular Lesions

Having concluded that the quantitative ribosome defect model best fit our data from complimentary human genetic and proteomic studies, we aimed to better understand the consequences of reduced ribosome abundances on translation on a global scale. To do so, we performed ribosome footprint profiling in the setting of RP haploinsuffciency in primary human HSPCs undergoing erythroid lineage commitment. This technique

involves measuring the translational efficiency (TE), by comparing the levels of ribosome-associated mRNA footprints to the total mRNA for each gene[25]. For biological replicates of RPL5 (uL18) and RPS19 (eS19) suppression, we obtained both ribosome-protected footprints (RPFs) and matching mRNA-sequencing (mRNA-seq). Changes in transcription and translation appeared to be largely similar between RPS19 and RPL5 haploinsufficiency (**Figure 4C**), which is consistent with the quantitative model, rather than a qualitative model in which ribosomes of distinct protein composition would regulate the translation of specific subsets of mRNAs.

At a false discovery rate (FDR) of 10%, we identified 525 transcripts whose TE was particularly sensitive to (and down-regulated by) RP haploinsufficiency and confirmed our previous finding that translation of *GATA1* mRNA is significantly decreased (**Figures 4D** and **4F**). We determined that this set of RP haploinsufficiency-sensitive transcripts is shorter in overall length, encoded more abundantly expressed proteins in unperturbed primary human erythroid progenitors[26], is translated more efficiently in physiologic conditions and is enriched for genes that are essential for erythroid cell growth and survival[27] (**Figures 4G** and **S4D**). Moreover, a subset of these transcripts is substantially up-regulated during early erythropoiesis[28], concurrent with the defective maturation stage in DBA, suggesting that the altered translation of multiple transcripts, including *GATA1*, plays a key role in the *in vivo* phenotypes observed in DBA. By comprehensively defining the actual 5' UTRs present in hematopoietic cells using cap analysis gene expression (CAGE) sequencing data (**Figure S5A**), we found that the 5' UTRs of RP haploinsufficiency-sensitive transcripts were 42 nucleotides shorter on average, were predicted to have less complex secondary structure, and contained fewer in-frame and out-of-frame upstream start codons (uAUGs) (**Figures 5A** and **S5C**) – features typically associated with efficient ribosome initiation in unperturbed cells[29].

4.5 Interrogation of 5' UTRs from Hematopoietic Master Regulators Suggests Mechanisms of Lineage Selectivity in DBA

While we could identify several mRNA- including 5' UTR-features that are associated with reduced mRNA translation due to ribosome defects caused by DBA-associated

molecular lesions, these findings are insufficient to explain the erythroid lineage-specific defect observed in DBA. Our genome-wide analysis has shown that translationally reduced mRNAs including *GATA1* have short and unstructured 5' UTRs (**Figure 5A**) and hence have properties to be efficiently translated under physiologic conditions[29]. As mentioned before, mutations in *GATA1* are sufficient to cause DBA in rare patients[16]. Thus, the reduced translation of *GATA1* mRNA could explain the perturbed erythroid lineage commitment in DBA. To understand why the production of other hematopoietic lineages is not affected in DBA, we examined CAGE data generated from unperturbed primary human HSPCs to compare the 5' UTRs among a group of 36 well-characterized hematopoietic master regulators known to have key and well-defined roles in lineage commitment[2,3] (**Figure 5B**). The majority of these other hematopoietic master regulators had significantly longer and more complex 5' UTRs compared to *GATA1* and compared with the other transcripts sensitive to reduced ribosome levels (**Figures 5C-5F**). In our ribosome profiling data, we were able to verify the lack of reduced translation in HSPCs undergoing erythroid lineage commitment with RP haploinsuffciency for master regulators that were sufficiently expressed in our cells: *KLF1*, *TAL1*, *MYB*, *GATA2*, *LMO2*, *RUNX2*, *ETV6*, *KMT2A*, *NFE2*, *FLI1*, *STAT5A*, *STAT3*, *SPI1*, *NOTCH1*, *BCL11A*, *IKZF1*, and *XBP1* all showed no major decrease in TE (FDR >10%, $\log_2$ TE fold decrease of <0.45). Taken together, these data suggest that *GATA1* exhibits relatively unique 5' UTR features among hematopoietic master regulators, which may explain its sensitivity to reduced ribosome levels and the consequent erythroid lineage-specific defect observed in DBA.

4.6 Impaired GATA1 Protein Production in Primary HSPCs from DBA Patients

We were able to show that cellular differentiation in erythropoiesis is regulated to a substantial extent by ribosome levels. A reduced number of ribosomes selectively perturbs the translation of a subset of mRNAs, among those is the key erythroid transcription factor *GATA1*, which appears to have relatively unique 5' UTR features among hematopoietic master regulators. This perturbation results in the erythroid

differentiation defects characteristic for DBA. We next aimed to identify the stages of hematopoiesis at which these impairments occur. By that we also aimed to assess what relevance the previously mentioned refined models of human hematopoiesis have with respect to human disease. These recent studies enabled by single cell technologies suggest that lineage commitment may occur already at early HSPC stages[4-6], that may be transcriptionally and epigenetically primed with biases towards certain lineages. Indeed, in the scope of these studies it was shown that *GATA1* mRNA is expressed already in primitive $CD34^+CD38^-$ HSPCs[4]. We were previously able to show that there is no difference in *GATA1* mRNA expression in early erythroid progenitors from DBA patients compared to healthy donors[10]. Here, we used an intracellular flow cytometry approach to assess GATA1 protein levels in $CD34^+CD38^-$ and $CD34^+CD38^+$ HSPCs from uncultured bone marrow aspirates from both DBA patients and healthy donors. For patients with *RPL35A*, *RPL5* or *RPS19* mutations, we consistently observed reduced GATA1 protein levels in $CD34^+CD38^-$ and $CD34^+CD38^+$ cell populations (**Figures 7B** and **7C**). As noted above, perturbations in GATA1 are sufficient to cause the erythroid phenotype characteristic for DBA[16]. Altogether, these findings are consistent with refined models of hematopoiesis, that suggest that lineage commitment occurs already in early HSPCs. Reduced GATA1 protein levels in this primitive cell compartment resulting from an impaired translation due to reduced ribosome levels cause erythroid differentiation defects typical for DBA.

4.7 Probing Alternative DBA Pathophysiological Models

To what extent alternative pathways, such as activation of p53 or heme excess, also contribute to the pathogenesis of DBA, is unclear, although our data suggest a major role for the translational defects as discussed in more detail in the discussion section (**please see under 5.3 and also Figures S8A-S8D (not shown in the original publication)**).

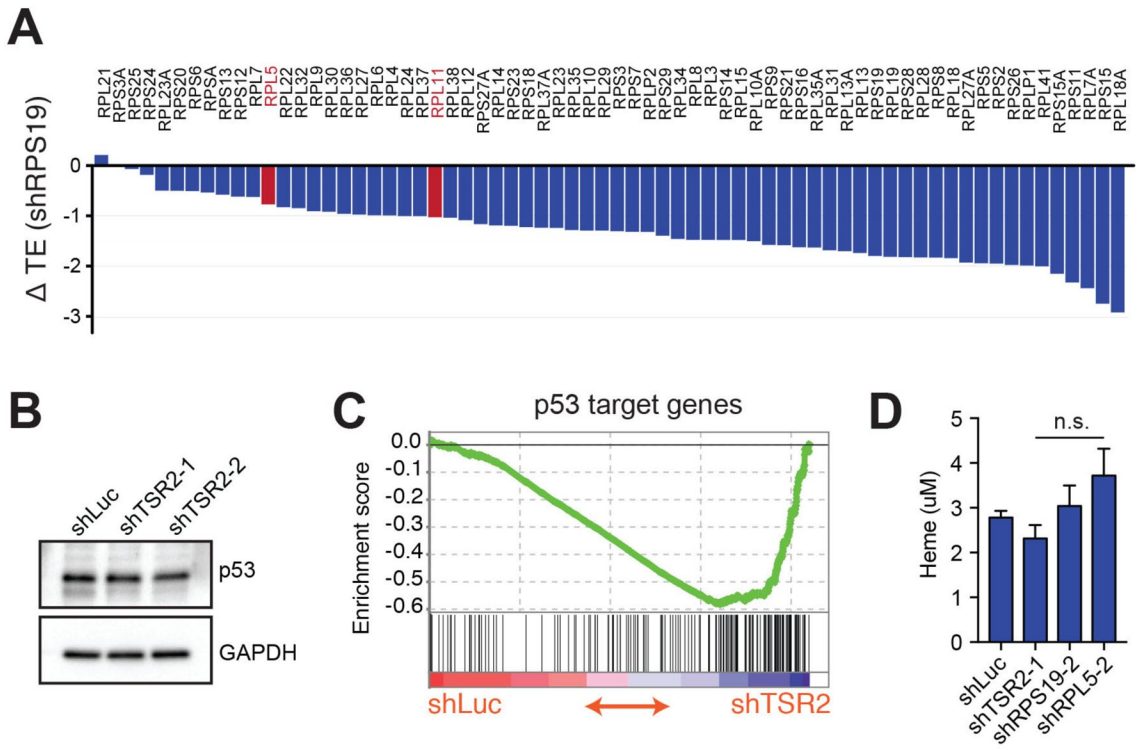


Figure S8 (not shown in the original publication). Limited evidence for alternative DBA pathophysiology models. (**A**) Log_2 changes in TE upon RPS19 suppression are shown. *RPL5* and *RPL11* are translationally co-regulated with other RPs. (**B**) Western blot for p53 protein levels upon TSR2 suppression in primary erythroid cells shows no evidence for substantial up-regulation of p53. (**C**) Based upon GSEA, cells with TSR2 suppression have a higher p53 target gene signature (permutation FDR < 0.0001). The enrichment score is plotted in green, and genes are plotted as black lines according to their rank. RNA-seq was performed in (n=2) biological replicates. (**D**) No evidence for significantly increased heme levels with ribosomal protein haploinsufficiency and TSR2 suppression models in primary erythroid cells.

5. Discussion

5.1 A Role for Ribosome Levels in Cellular Differentiation in Human Hematopoiesis

Recent studies have tremendously improved our understanding of hematopoiesis and suggest that lineage commitment occurs already at early HSPC stages[4-6]. The relevance of these refined models of hematopoiesis to human disease has remained unclear. Furthermore, the vast majority of these studies has focused on the regulation of these processes on the transcriptional level[4,5], while the role of post-transcriptional regulation including of the molecular key regulators of lineage commitment has been largely unexplored. Here, we used the rare congenital blood disorder DBA to address these issues to some extent. The majority of DBA cases are caused by mutations in one of at least 26 different ribosome protein genes, resulting in ribosomal protein haploinsufficiency[15]. We find that in case of DBA, reduced ribosome levels, rather than qualitatively distinct ribosomes, result in the impaired translation of a subset of mRNAs. These transcripts tend to have shorter and less structured 5' UTRs, features associated with efficient translation in unperturbed conditions[29]. Among those transcripts is the key erythroid transcription factor GATA1, which exhibits these features relatively uniquely among hematopoietic master regulators, which may explain the erythroid specific defect seen in DBA within the hematopoietic compartment. Finally, we are able to show that GATA1 protein levels are already reduced in HSPCs from bone marrow aspirates from DBA patients, consistent with revised models of hematopoiesis, where lineage commitment occurs in this primitive cell compartment[4-6]. Our work further complements studies by others showing substantial variation in protein synthesis rates in hematopoiesis, with the highest protein synthesis rates found in progenitors that undergo erythroid lineage commitment[8], which fits to our data of how ribosome levels can selectively impair erythroid lineage commitment and *GATA1* requires one of the highest translation rates among various master regulators of hematopoiesis. Given that GATA1 is a common downstream target of different DBA-associated molecular lesions, it presents a promising potential target for clinical therapy.

The role of post-transcriptional regulation in development, tissue homeostasis and cellular differentiation is increasingly appreciated[30-32], but still largely unexplored. An emerging role in this regard is attributed to the ribosome. Indeed, the so called ribosomopathies represent a group of diseases with a wide range of phenotypes such as congenital asplenia, neurodevelopmental disorders[12] and DBA including its non-hematopoietic phenotypes such as cleft lip/ palate and thumb abnormalities, all of which are caused by diverse perturbations in the ribosome[33]. How such tissue-specific defects are mediated by lesions in components of a machinery so ubiquitous as the ribosome, is subject to current, and some recent investigations including this work. While in some contexts altered RP composition has been suggested to underlie altered translation of subsets of mRNAs and thereof resulting tissue-specific patterning defects[18], in case of DBA, we find a constant RP composition, but reduced levels of ribosomes that selectively perturb the translation of a subset of transcripts, thereby impeding erythroid lineage commitment. Future studies will be required to delineate if such mechanisms also underlie other phenotypes that are caused by perturbations in the ribosome. Curiously, variation in ribosome levels also seems to underlie phenotypic changes within a single cell type, suggesting that such regulation of ribosome levels may also alter cell state in some contexts[34]. Here, we discovered an unprecedented role for ribosome levels in cellular differentiation in human hematopoiesis and translation regulation.

5.2 Discriminating between Ribosome Defect Models

In this study, we have investigated if the actively translating ribosomes in hematopoietic cells with ribosomal protein haploinsufficiency or with TSR2 suppression lack the targeted RP (RPL5 (uL18), RPS19 (eS19), or RPS26 (eS26)) – what we term the "qualitative" model. We first showed that RP intensities were highly correlated in LPs and HPs. Visually, the targeted RP seemed to fall close to the line determined from both total and targeted subunit specific linear regression. To make a more quantitative argument, we investigated exactly how well this line predicted the intensity of each RP by calculating studentized residuals for each fit. In general, studentized residuals were small (values greater than 3 or less than -3 are usually considered outliers). As the

targeted RP was not observed at the negative tail of the distribution of these residuals, we determined that it was highly unlikely that ribosomes in LPs or HPs were lacking a single, targeted RP. Thus, we concluded that our data best fit the "quantitative" model where DBA-associated molecular lesions result in reduced levels of ribosomes where each ribosome has identical RP composition and contains the targeted RP.

It is not known, however, if each ribosome is identically composed. For example, it is possible that there are multiple types of ribosomes within a cell or given fraction. In this hypothetical case, each type of ribosome is composed of a unique set of RPs. This possibility is supported by recent evidence showing that RP stoichiometry can vary between monosome and polysome fractions[35], although it remains unclear if there are indeed differently composed ribosomes within a single fraction. Hypothetically, as illustrated in a simplified manner in **Figure S9**, *type A* ribosomes could contain only RP1, RP2, RP3, RP4, and RP5, *type B* ribosomes could contain only RP1, RP2, RP4, RP6, and

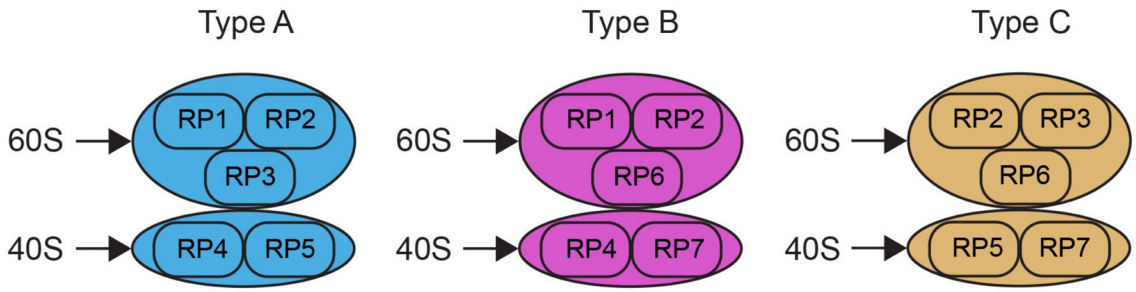


Figure S9 (not shown in the original publication). A simplified scheme of potential different types of ribosomes within a given fraction. Given that we are measuring the average RP intensity across all ribosomes within the M, LP, and HP fractions, we cannot exclude the possibility that there might be different types of ribosomes within a given fraction and cannot exclude the possibility of changes in the relative abundances of these types with ribosomal protein haploinsufficiency.

RP7, and *type C* ribosomes could contain only RP2, RP3, RP5, RP6, and RP7. Given that we are measuring the average RP intensity across all ribosomes within the M, LP, and HP fractions, we cannot determine the composition of any individual ribosome and thus do not know (1) if there are differently composed ribosomes, (2) the number of ribosome types, or (3) the relative abundances of these types. In case that there is only one type, our data strongly support that ribosome composition is unaltered with the DBA-associated molecular lesions tested. In case that there are a number of ribosome types capable of mRNA translation, then our data again strongly support the situation where each

ribosome type is equivalently reduced with ribosomal protein haploinsufficiency or TSR2 suppression, as (1) RP intensities are very highly correlated between DBA-associated molecular lesions and control conditions and (2) no RPs are identified as significant outliers (residual less than -3). However, we cannot definitively exclude the possibility that the relative weights of these potential different types of ribosomes are altered with DBA-associated molecular lesions. More formally, assuming m ribosome types, the observed intensity for the ith RP (RP_i) is proportional to $\sum_{j \in m} W_j I_{ji}$ where W_j is the percentage of ribosomes of type j and I_{ji} is an indicator variable for whether RP_i is present in type j. In this scenario, although the relative proportion of ribosome types changes (i.e. altered W_j), the observed intensity for each RP_i (i.e. unchanged $\sum_{j \in m} W_j I_{ji}$) is unchanged – at least within the technical measurement error. As this would require very specific biological types of ribosomes and a very specific alteration in the proportion of these types, this seems unlikely but remains a possibility. Of note in this context, RPs implicated in DBA appear to be scattered all over the ribosomal surface based on structural studies of the ribosome[15,36]. Studies of individual (types of) ribosomes will help to further clarify this point. A comprehensive overview about the emerging field of ribosome heterogeneity can be found in a perspective by Emmott and colleagues[17].

5.3 Alternative DBA Pathophysiological Models

One prevalent theory for the erythroid specific phenotype of DBA is that p53 is selectively activated in the erythroid lineage[37-41]. In the case of ribosomal protein haploinsufficiency due to *RPS19* mutations, it has been proposed that excess free RPL5 (uL18) and RPL11 (uL5) would bind to and inhibit mouse double minute 2 (MDM2)[37,39,41-43], a ubiquitin ligase that regulates the stability of p53 by targeting its degradation[41,44]. This in turn leads to the accumulation of p53 and consequently to cell cycle arrest and apoptosis[40,41]. In agreement with this theory, disruption of the RPL5- or RPL11-MDM2 interaction by a point mutation in *Mdm2* was shown to reverse the p53 response and lead to an amelioration of the anemia in a RPS19 deficient mouse model of DBA[41,45,46]. On the other hand, a number of recent studies have suggested that the erythroid defect occurs independent of p53 in both mouse and zebrafish models of DBA,

challenging the theory that the lineage specific defect is mediated by RP-mediated MDM2 inhibition and p53 activation[47-51].

To what extent such a mechanism contributes to the erythroid specific phenotype in DBA, remains controversial[41]. For instance, while RPL5 and RPL11 have been demonstrated to be crucial for MDM2-mediated p53 activation, both genes are frequently mutated in DBA, ultimately resulting in decreased (rather than increased) free RPL5 or RPL11[15,43]. Here, in our RPS19 haploinsufficent model of DBA in primary human erythroid cells, we observe that RPL5 and RPL11 are, similar to all other RPs, translationally co-down-regulated (**Figure S8A**), suggesting that there is a limited amount of free RPL5 or RPL11 capable of binding MDM2. Similar to what we previously observed with ribosomal protein haploinsufficiency[10], p53 protein levels do not appear to be substantially increased in our models of TSR2 suppression (**Figure S8B**). When we previously interrogated erythroid cells isolated from DBA patients, we did not detect a global up-regulation of p53 target genes[10]. Furthermore, we recently reported that the inhibition of the eIF4E and eIF4G interaction by 4EGI-1 phenocopied the typical erythroid defect observed in DBA, but did not result in ribosomal defects or an imbalance of RPs[10,41]. Together, these data suggest that the erythroid defect occurs largely independent from RP-mediated p53 activation.

Interestingly, GATA1 has been suggested to associate with p53 and potentially inhibit its pro-apoptotic activities[40,41,52]. In agreement with these studies, we previously demonstrated that suppression of GATA1 results in increased p53 levels in human primary erythroid cells[10,41]. However, GATA1 has been shown to repress pro-apoptotic and activate anti-apoptotic genes independent from its interaction with p53[41,53,54]. In our ribosomal protein haploinsufficiency and TSR2 suppression models, we do indeed observe a relative, but not global increase in p53 target genes, many of which are pro-apoptotic[10,41] (**Figure S8C**). Thus, the impaired translation of *GATA1* – and potentially other selectively impaired transcripts – may result in apoptosis through multiple p53 dependent and independent mechanisms[41]. Further work is required to delineate the exact contribution of p53 activation to the lineage specific defect in DBA, but current evidence suggests that RP-mediated (through MDM2 inhibition) activation of p53 is a relatively limited factor.

In addition to the model of RP-mediated p53 activation, a second DBA pathophysiological model hypothesized that the macrocytic anemia in DBA could be the consequence of a toxic heme excess that results from delayed globin synthesis[55]. In this DBA model, heme synthesis proceeds normally but globin production is delayed, resulting in an excess of free heme. Since this potential theory is just emerging and the data supporting it are still limited, we examined heme levels in our human primary erythroid cell models of DBA due to ribosomal protein haploinsufficiency or TSR2 suppression, but did not observe significantly increased heme levels (**Figure S8D**). A subsequent study of the same group used Flvcr1-deleted mice, whereby this deletion leads to high intracellular heme levels as Flvcr1 is a cytoplasmic heme export protein, to further pursue this hypothesis[56]. The authors suggest a model of a heme-GATA1 feedback loop, where GATA1 stimulates heme synthesis and heme would decrease GATA1 levels in the course of erythroid differentiation in physiologic conditions. This decrease of GATA1 not only included the protein, but also the *GATA1* transcript. The authors further speculate, that a pathologic increase of heme levels could result in premature reduction of GATA1 thereby impeding erythroid differentiation and resulting in the phenotype typical for DBA[56]. However, we previously demonstrated that there is no difference in *GATA1* mRNA expression in early erythroid progenitors from DBA patients compared to healthy donors[10]. Given these findings and that our human primary erythroid cell models of DBA phenocopy the erythroid specific defect of this disease and we don't find significantly increased heme levels in these, suggests that the mouse model used by the authors might not ideally resemble the phenotype seen in DBA patients and that heme excess has a limited role in the pathogenesis of DBA, but more direct studies are needed to quantify the exact contribution of this potential mechanism to the clinical phenotype.

6. Bibliography

1. Khajuria RK, Munschauer M, Ulirsch JC, Fiorini C, Ludwig LS, McFarland SK, Abdulhay NJ, Specht H, Keshishian H, Mani DR, Jovanovic M, Ellis SR, Fulco CP, Engreitz JM, Schutz S, Lian J, Gripp KW, Weinberg OK, Pinkus GS, Gehrke L, Regev A, Lander ES, Gazda HT, Lee WY, Panse VG, Carr SA, Sankaran VG. Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis. Cell 2018;173:90-103 e19.
2. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. Cell 2008;132:631-44.
3. Doulatov S, Notta F, Laurenti E, Dick JE. Hematopoiesis: a human perspective. Cell stem cell 2012;10:120-36.
4. Notta F, Zandi S, Takayama N, Dobson S, Gan OI, Wilson G, Kaufmann KB, McLeod J, Laurenti E, Dunant CF, McPherson JD, Stein LD, Dror Y, Dick JE. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science 2016;351:aab2116.
5. Paul F, Arkin Y, Giladi A, Jaitin DA, Kenigsberg E, Keren-Shaul H, Winter D, Lara-Astiaso D, Gury M, Weiner A, David E, Cohen N, Lauridsen FK, Haas S, Schlitzer A, Mildner A, Ginhoux F, Jung S, Trumpp A, Porse BT, Tanay A, Amit I. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. Cell 2015;163:1663-77.
6. Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, Hirche C, Lutz C, Buss EC, Nowak D, Boch T, Hofmann WK, Ho AD, Huber W, Trumpp A, Essers MA, Steinmetz LM. Human haematopoietic stem cell lineage commitment is a continuous process. Nat Cell Biol 2017;19:271-81.
7. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. Global quantification of mammalian gene expression control. Nature 2011;473:337-42.
8. Signer RA, Magee JA, Salic A, Morrison SJ. Haematopoietic stem cells require a highly regulated protein synthesis rate. Nature 2014;509:49-54.
9. McCann KL, Baserga SJ. Genetics. Mysterious ribosomopathies. Science 2013;341:849-50.
10. Ludwig LS, Gazda HT, Eng JC, Eichhorn SW, Thiru P, Ghazvinian R, George TI, Gotlib JR, Beggs AH, Sieff CA, Lodish HF, Lander ES, Sankaran VG. Altered translation of GATA1 in Diamond-Blackfan anemia. Nat Med 2014;20:748-53.
11. Bolze A, Mahlaoui N, Byun M, Turner B, Trede N, Ellis SR, Abhyankar A, Itan Y, Patin E, Brebner S, Sackstein P, Puel A, Picard C, Abel L, Quintana-Murci L, Faust SN, Williams AP, Baretto R, Duddridge M, Kini U, Pollard AJ, Gaud C, Frange P, Orbach D, Emile JF, Stephan JL, Sorensen R, Plebani A, Hammarstrom L, Conley ME, Selleri L, Casanova JL. Ribosomal protein SA haploinsufficiency in humans with isolated congenital asplenia. Science 2013;340:976-8.
12. Brooks SS, Wall AL, Golzio C, Reid DW, Kondyles A, Willer JR, Botti C, Nicchitta CV, Katsanis N, Davis EE. A novel ribosomopathy caused by dysfunction of RPL10 disrupts neurodevelopment and causes X-linked microcephaly in humans. Genetics 2014;198:723-33.
13. Nathan DG, Clarke BJ, Hillman DG, Alter BP, Housman DE. Erythroid precursors in congenital hypoplastic (Diamond-Blackfan) anemia. The Journal of clinical investigation 1978;61:489-98.
14. Iskander D, Psaila B, Gerrard G, Chaidos A, En Foong H, Harrington Y, Karnik LC, Roberts I, de la Fuente J, Karadimitris A. Elucidation of the EP defect in Diamond-Blackfan anemia by characterization and prospective isolation of human EPs. Blood 2015;125:2553-7.
15. Ulirsch JC, Verboon JM, Kazerounian S, Guo MH, Yuan D, Ludwig LS, Handsaker RE, Abdulhay NJ, Fiorini C, Genovese G, Lim ET, Cheng A, Cummings BB, Chao KR, Beggs AH, Genetti CA, Sieff CA, Newburger PE, Niewiadomska E, Matysiak M, Vlachos A, Lipton JM, Atsidaftos E, Glader B, Narla A, Gleizes PE, O'Donohue MF, Montel-Lehry N, Amor DJ, McCarroll SA, O'Donnell-Luria AH, Gupta N, Gabriel SB, MacArthur DG, Lander ES, Lek M, Da Costa L, Nathan DG, Korostelev AA, Do R, Sankaran VG, Gazda HT. The Genetic Landscape of Diamond-Blackfan Anemia. Am J Hum Genet 2018;103:930-47.
16. Sankaran VG, Ghazvinian R, Do R, Thiru P, Vergilio JA, Beggs AH, Sieff CA, Orkin SH, Nathan DG, Lander ES, Gazda HT. Exome sequencing identifies GATA1 mutations resulting in Diamond-Blackfan anemia. The Journal of clinical investigation 2012;122:2439-43.

17. Emmott E, Jovanovic M, Slavov N. Ribosome Stoichiometry: From Form to Function. Trends Biochem Sci 2019;44:95-109.
18. Kondrashov N, Pusic A, Stumpf CR, Shimizu K, Hsieh AC, Ishijima J, Shiroishi T, Barna M. Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. Cell 2011;145:383-97.
19. Lodish HF. Model for the regulation of mRNA translation applied to haemoglobin synthesis. Nature 1974;251:385-8.
20. Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES, Plath K, Guttman M. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science 2013;341:1237973.
21. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545-50.
22. Gripp KW, Curry C, Olney AH, Sandoval C, Fisher J, Chong JX, Genomics UWCfM, Pilchman L, Sahraoui R, Stabley DL, Sol-Church K. Diamond-Blackfan anemia with mandibulofacial dystostosis is heterogeneous, including the novel DBA genes TSR2 and RPS28. Am J Med Genet A 2014;164A:2240-9.
23. Schutz S, Fischer U, Altvater M, Nerurkar P, Pena C, Gerber M, Chang Y, Caesar S, Schubert OT, Schlenstedt G, Panse VG. A RanGTP-independent mechanism allows ribosomal protein nuclear import for ribosome assembly. Elife 2014;3:e03473.
24. Quarello P, Garelli E, Carando A, Mancini C, Foglia L, Botto C, Farruggia P, De Keersmaecker K, Aspesi A, Ellis SR, Dianzani I, Ramenghi U. Ribosomal RNA analysis in the diagnosis of Diamond-Blackfan Anaemia. British journal of haematology 2016;172:782-5.
25. Ingolia NT. Ribosome Footprint Profiling of Translation throughout the Genome. Cell 2016;165:22-33.
26. Gautier EF, Ducamp S, Leduc M, Salnot V, Guillonneau F, Dussiot M, Hale J, Giarratana MC, Raimbault A, Douay L, Lacombe C, Mohandas N, Verdier F, Zermati Y, Mayeux P. Comprehensive Proteomic Analysis of Human Erythropoiesis. Cell reports 2016;16:1470-84.
27. Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM. Identification and characterization of essential genes in the human genome. Science 2015;350:1096-101.
28. Li J, Hale J, Bhagia P, Xue F, Chen L, Jaffray J, Yan H, Lane J, Gallagher PG, Mohandas N, Liu J, An X. Isolation and transcriptome analyses of human erythroid progenitors: BFU-E and CFU-E. Blood 2014;124:3636-45.
29. Hinnebusch AG, Ivanov IP, Sonenberg N. Translational control by 5'-untranslated regions of eukaryotic mRNAs. Science 2016;352:1413-6.
30. Buszczak M, Signer RA, Morrison SJ. Cellular differences in protein synthesis regulate tissue homeostasis. Cell 2014;159:242-51.
31. Baser A, Skabkin M, Kleber S, Dang Y, Gulculer Balta GS, Kalamakis G, Gopferich M, Ibanez DC, Schefzik R, Lopez AS, Bobadilla EL, Schultz C, Fischer B, Martin-Villalba A. Onset of differentiation is post-transcriptionally controlled in adult neural stem cells. Nature 2019;566:100-4.
32. Chau KF, Shannon ML, Fame RM, Fonseca E, Mullan H, Johnson MB, Sendamarai AK, Springel MW, Laurent B, Lehtinen MK. Downregulation of ribosome biogenesis during early forebrain development. Elife 2018;7.
33. Danilova N, Gazda HT. Ribosomopathies: how a common root can cause a tree of pathologies. Dis Model Mech 2015;8:1013-26.
34. Sinturel F, Gerber A, Mauvoisin D, Wang J, Gatfield D, Stubblefield JJ, Green CB, Gachon F, Schibler U. Diurnal Oscillations in Liver Mass and Cell Size Accompany Ribosome Assembly Cycles. Cell 2017;169:651-63 e14.
35. Slavov N, Semrau S, Airoldi E, Budnik B, van Oudenaarden A. Differential Stoichiometry among Core Ribosomal Proteins. Cell reports 2015;13:865-73.
36. Ben-Shem A, Garreau de Loubresse N, Melnikov S, Jenner L, Yusupova G, Yusupov M. The structure of the eukaryotic ribosome at 3.0 A resolution. Science 2011;334:1524-9.

37. Narla A, Ebert BL. Ribosomopathies: human disorders of ribosome dysfunction. Blood 2010;115:3196-205.
38. Dutt S, Narla A, Lin K, Mullally A, Abayasekara N, Megerdichian C, Wilson FH, Currie T, Khanna-Gupta A, Berliner N, Kutok JL, Ebert BL. Haploinsufficiency for ribosomal protein genes causes selective activation of p53 in human erythroid progenitor cells. Blood 2011;117:2567-76.
39. Zhang Y, Lu H. Signaling to p53: ribosomal proteins find their way. Cancer cell 2009;16:369-77.
40. Weiss MJ, Mason PJ, Bessler M. What's in a name? The Journal of clinical investigation 2012;122:2346-9.
41. Ludwig LS-H. Altered translation of GATA1 in Diamond Blackfan Anemia and the role of cyclin D3 in erythropoiesis [Berlin, Freie Univ., Diss., 2015]2015.
42. Marechal V, Elenbaas B, Piette J, Nicolas JC, Levine AJ. The ribosomal L5 protein is associated with mdm-2 and mdm-2-p53 complexes. Molecular and cellular biology 1994;14:7414-20.
43. Bursac S, Brdovcak MC, Pfannkuchen M, Orsolic I, Golomb L, Zhu Y, Katz C, Daftuar L, Grabusic K, Vukelic I, Filic V, Oren M, Prives C, Volarevic S. Mutual protection of ribosomal proteins L5 and L11 from degradation is essential for p53 activation upon ribosomal biogenesis stress. Proc Natl Acad Sci U S A 2012;109:20467-72.
44. Fang S, Jensen JP, Ludwig RL, Vousden KH, Weissman AM. Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself and p53. The Journal of biological chemistry 2000;275:8945-51.
45. Macias E, Jin A, Deisenroth C, Bhat K, Mao H, Lindstrom MS, Zhang Y. An ARF-independent c-MYC-activated tumor suppression pathway mediated by ribosomal protein-Mdm2 Interaction. Cancer cell 2010;18:231-43.
46. Jaako P, Debnath S, Olsson K, Zhang Y, Flygare J, Lindstrom MS, Bryder D, Karlsson S. Disruption of the 5S RNP-Mdm2 interaction significantly improves the erythroid defect in a mouse model for Diamond-Blackfan anemia. Leukemia 2015;29:2221-9.
47. Torihara H, Uechi T, Chakraborty A, Shinya M, Sakai N, Kenmochi N. Erythropoiesis failure due to RPS19 deficiency is independent of an activated Tp53 response in a zebrafish model of Diamond-Blackfan anaemia. British journal of haematology 2011;152:648-54.
48. Yadav GV, Chakraborty A, Uechi T, Kenmochi N. Ribosomal protein deficiency causes Tp53-independent erythropoiesis failure in zebrafish. The international journal of biochemistry & cell biology 2014;49:1-7.
49. Singh SA, Goldberg TA, Henson AL, Husain-Krautter S, Nihrane A, Blanc L, Ellis SR, Lipton JM, Liu JM. p53-Independent cell cycle and erythroid differentiation defects in murine embryonic stem cells haploinsufficient for Diamond Blackfan anemia-proteins: RPS19 versus RPL5. PloS one 2014;9:e89098.
50. Antunes AT, Goos YJ, Pereboom TC, Hermkens D, Wlodarski MW, Da Costa L, MacInnes AW. Ribosomal Protein Mutations Result in Constitutive p53 Protein Degradation through Impairment of the AKT Pathway. PLoS Genet 2015;11:e1005326.
51. Morgado-Palacin L, Varetti G, Llanos S, Gomez-Lopez G, Martinez D, Serrano M. Partial Loss of Rpl11 in Adult Mice Recapitulates Diamond-Blackfan Anemia and Promotes Lymphomagenesis. Cell reports 2015;13:712-22.
52. Trainor CD, Mas C, Archambault P, Di Lello P, Omichinski JG. GATA-1 associates with and inhibits p53. Blood 2009;114:165-73.
53. Maeda T, Ito K, Merghoub T, Poliseno L, Hobbs RM, Wang G, Dong L, Maeda M, Dore LC, Zelent A, Luzzatto L, Teruya-Feldstein J, Weiss MJ, Pandolfi PP. LRF is an essential downstream target of GATA1 in erythroid development and regulates BIM-dependent apoptosis. Developmental cell 2009;17:527-40.
54. Cheng Y, Wu W, Kumar SA, Yu D, Deng W, Tripic T, King DC, Chen KB, Zhang Y, Drautz D, Giardine B, Schuster SC, Miller W, Chiaromonte F, Zhang Y, Blobel GA, Weiss MJ, Hardison RC. Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. Genome research 2009;19:2172-84.
55. Yang Z, Keel SB, Shimamura A, Liu L, Gerds AT, Li HY, Wood BL, Scott BL, Abkowitz JL. Delayed globin synthesis leads to excess heme and the macrocytic anemia of Diamond Blackfan anemia and del(5q) myelodysplastic syndrome. Science translational medicine 2016;8:338ra67.
56. Doty RT, Yan X, Lausted C, Munday AD, Yang Z, Yi D, Jabbari N, Liu L, Keel SB, Tian Q, Abkowitz JL. Single-cell analyses demonstrate that a heme-GATA1 feedback loop regulates red cell differentiation. Blood 2019;133:457-69.

7. Affidavit/ Eidesstattliche Versicherung

„Ich, Rajiv Kumar Khajuria, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Consequences of Ribosomal Perturbations in Human Hematopoiesis" selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.
Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen werden von mir verantwortet.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Betreuer/in, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; www.icmje.og) zur Autorenschaft eingehalten. Ich erkläre ferner, dass mir die Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis bekannt ist und ich mich zur Einhaltung dieser Satzung verpflichte.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst."

Datum: 18.06.2020 _____ Unterschrift

8. Declaration of Contribution/ Anteilserklärung

Publication:

Khajuria RK, Munschauer M*, Ulirsch JC*, Fiorini C*, Ludwig LS*, McFarland SK, Abdulhay NJ, Specht H, Keshishian H, Mani DR, Jovanovic M, Ellis SR, Fulco CP, Engreitz JM, Schütz S, Lian J, Gripp KW, Weinberg OK, Pinkus GS, Gehrke L, Regev A, Lander ES, Gazda HT, Lee WY, Panse VG, Carr SA, Sankaran VG. *Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis.* ***Cell*** *2018 Mar 22;173(1):90-103.e19. doi: 10.1016/j.cell.2018.02.036. Epub 2018 Mar 15*

(*equal contribution)

Contribution / Beitrag im Einzelnen:

Rajiv Kumar Khajuria designed the study with Vijay G. Sankaran. He designed, performed and analyzed most experiments in the study including in vitro culture of hematopoietic stem and progenitor and erythroid cells, mass spectrometry experiments, ribosome profiling, polysome profiling, lentiviral infections, flow cytometric assays, RT-qPCRs, western blot experiments and coordinated bioinformatic analyses with Jacob C. Ulirsch, D.R. Mani and Harrison Specht. He wrote the manuscript with Jacob C. Ulirsch and Vijay G. Sankaran with input from all authors. More specifically, he was involved in designing the experiments, analyzing/visualizing and/or acquiring the data for the following figures: Figures 1A and 1C-H, Figures 2B-G, Figures 3A-D, Figures 4A-G, Figures 5A-F, Figures 7B-C, Figures S1B-L, Figures S2A-Q, Figures S3A-K, Figures S4A-D and Figures S5B-F.

_____

Unterschrift, Datum und Stempel der betreuenden Hochschullehrerin

_____

Unterschrift des Doktoranden

9. Excerpt from the Journal Summary List

Journal Data Filtered By: **Selected JCR Year: 2017** Selected Editions: SCIE,SSCI
Selected Categories: **"CELL BIOLOGY"** Selected Category Scheme: WoS
Gesamtanzahl: 190 Journale

Rank	Full Journal Title	Total Cites	Journal Impact Factor	Eigenfactor Score
1	NATURE REVIEWS MOLECULAR CELL BIOLOGY	43,667	35.612	0.095540
2	NATURE MEDICINE	75,461	32.621	0.171980
3	**CELL**	**230,625**	**31.398**	**0.583260**
4	Cell Stem Cell	23,493	23.290	0.096030
5	CANCER CELL	35,217	22.844	0.096910
6	Cell Metabolism	29,834	20.565	0.101740
7	NATURE CELL BIOLOGY	39,896	19.064	0.092960
8	TRENDS IN CELL BIOLOGY	13,708	18.564	0.037630
9	Science Translational Medicine	26,691	16.710	0.126450
10	CELL RESEARCH	13,728	15.393	0.037450
11	MOLECULAR CELL	61,604	14.248	0.181170
12	NATURE STRUCTURAL & MOLECULAR BIOLOGY	27,547	13.333	0.081820
13	Autophagy	14,923	11.100	0.035510
14	TRENDS IN MOLECULAR MEDICINE	9,213	11.021	0.019720
15	EMBO JOURNAL	67,036	10.557	0.079780
16	CURRENT OPINION IN CELL BIOLOGY	13,339	10.015	0.027790
17	DEVELOPMENTAL CELL	26,896	9.616	0.074980
18	GENES & DEVELOPMENT	57,469	9.462	0.092720
19	CURRENT BIOLOGY	56,595	9.251	0.137200
20	Cold Spring Harbor Perspectives in Biology	13,275	9.247	0.049360

Selected JCR Year: 2017; Selected Categories: "CELL BIOLOGY"

10. Publication

Khajuria RK, Munschauer M*, Ulirsch JC*, Fiorini C*, Ludwig LS*, McFarland SK, Abdulhay NJ, Specht H, Keshishian H, Mani DR, Jovanovic M, Ellis SR, Fulco CP, Engreitz JM, Schütz S, Lian J, Gripp KW, Weinberg OK, Pinkus GS, Gehrke L, Regev A, Lander ES, Gazda HT, Lee WY, Panse VG, Carr SA, Sankaran VG. *Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis.* ***Cell*** *2018 Mar 22;173(1):90-103.e19. doi: 10.1016/j.cell.2018.02.036. Epub 2018 Mar 15*

(*equal contribution)

Impact factor (in year of publication): 36.216

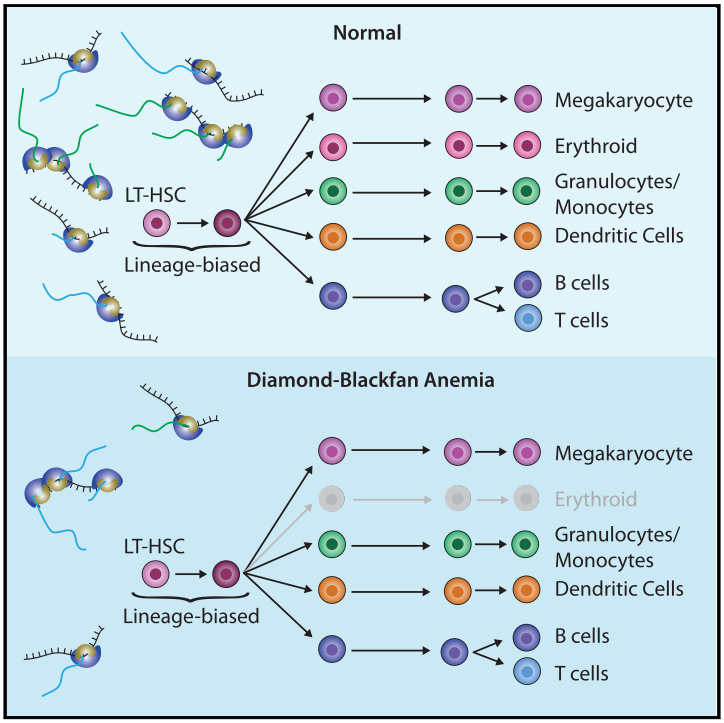
DOI: https://doi.org/10.1016/j.cell.2018.02.036

Article

Cell

Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis

Graphical Abstract



Highlights

- Molecular lesions underlying DBA reduce ribosome levels in hematopoietic cells
- Ribosome composition remains constant in cells with DBA-associated lesions
- Reduced ribosome levels selectively impair translation of a subset of mRNAs
- Translational perturbations in DBA impair lineage-commitment in HSPCs

Authors

Rajiv K. Khajuria, Mathias Munschauer, Jacob C. Ulirsch, ..., Vikram G. Panse, Steven A. Carr, Vijay G. Sankaran

Correspondence

sankaran@broadinstitute.org

In Brief

A global reduction in ribosome levels in Diamond-Blackfan anemia profoundly alters translation of a select subset of transcripts, thereby impeding erythroid lineage commitment.

Khajuria et al., 2018, Cell *173*, 90–103
March 22, 2018
https://doi.org/10.1016/j.cell.2018.02.036

CellPress

Cell

Article

Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis

Rajiv K. Khajuria,[1,2,3] Mathias Munschauer,[2,13] Jacob C. Ulirsch,[1,2,13] Claudia Fiorini,[1,2,13] Leif S. Ludwig,[1,2,13] Sean K. McFarland,[1,2] Nour J. Abdulhay,[1,2] Harrison Specht,[2] Hasmik Keshishian,[2] D.R. Mani,[2] Marko Jovanovic,[2] Steven R. Ellis,[4] Charles P. Fulco,[2] Jesse M. Engreitz,[2] Sabina Schütz,[5] John Lian,[6,7,8] Karen W. Gripp,[9] Olga K. Weinberg,[10] Geraldine S. Pinkus,[11] Lee Gehrke,[6,7,8] Aviv Regev,[2] Eric S. Lander,[2] Hanna T. Gazda,[2,12] Winston Y. Lee,[11] Vikram G. Panse,[5] Steven A. Carr,[2] and Vijay G. Sankaran[1,2,14,*]

[1]Division of Hematology/Oncology, Boston Children's Hospital and Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA
[2]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[3]Berlin-Brandenburg School for Regenerative Therapies, Charité-Universitätsmedizin Berlin, Berlin 13353, Germany
[4]Department of Biochemistry and Molecular Biology, University of Louisville, Louisville, KY 40202, USA
[5]Institute of Medical Microbiology, Department of Medicine, University of Zurich, 8006 Zurich, Switzerland
[6]Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[7]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115, USA
[8]Harvard-MIT Program in Health Sciences and Technology, Cambridge, MA 02139, USA
[9]Division of Medical Genetics, A. I. duPont Hospital for Children, Wilmington, DE 19803, USA
[10]Department of Pathology, Boston Children's Hospital, Boston, MA 02115, USA
[11]Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA
[12]Division of Genetics and Genomics, Manton Center for Orphan Disease Research, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA
[13]These authors contributed equally
[14]Lead Contact
*Correspondence: sankaran@broadinstitute.org
https://doi.org/10.1016/j.cell.2018.02.036

SUMMARY

Blood cell formation is classically thought to occur through a hierarchical differentiation process, although recent studies have shown that lineage commitment may occur earlier in hematopoietic stem and progenitor cells (HSPCs). The relevance to human blood diseases and the underlying regulation of these refined models remain poorly understood. By studying a genetic blood disorder, Diamond-Blackfan anemia (DBA), where the majority of mutations affect ribosomal proteins and the erythroid lineage is selectively perturbed, we are able to gain mechanistic insight into how lineage commitment is programmed normally and disrupted in disease. We show that in DBA, the pool of available ribosomes is limited, while ribosome composition remains constant. Surprisingly, this global reduction in ribosome levels more profoundly alters translation of a select subset of transcripts. We show how the reduced translation of select transcripts in HSPCs can impair erythroid lineage commitment, illuminating a regulatory role for ribosome levels in cellular differentiation.

INTRODUCTION

Blood cell production or hematopoiesis serves as a paradigm for cellular differentiation more generally in physiologic systems (Orkin and Zon, 2008). Extensive work has revealed a hierarchical progression of differentiation, where increasingly more lineage-restricted progenitors are produced, ultimately giving rise to lineage committed progenitors and precursors that eventually form mature circulating blood cells (Doulatov et al., 2012; Orkin and Zon, 2008). These observations have served as a framework for understanding the molecular regulation of hematopoiesis and how this process can be perturbed in disease. However, the majority of studies characterizing hematopoiesis in humans and mice have required analysis of bulk cell populations. Recent work, enabled through single-cell analyses and refined phenotypic markers, has shown that hematopoietic differentiation may progress in a distinct manner, where lineage commitment occurs in early hematopoietic stem and progenitor cells (HSPCs) that then undergo orderly differentiation to produce mature circulating blood cells (Notta et al., 2016; Paul et al., 2015; Perié et al., 2015; Velten et al., 2017).

While considerable insight into lineage commitment from HSPCs has been gained at the transcriptional level (Notta et al., 2016; Paul et al., 2015), the repertoire of molecular regulators of this process remains to be fully defined and functionally characterized. Groundbreaking studies have revealed the key role of post-transcriptional regulation in the maintenance of

hematopoietic stem cells (Signer et al., 2014; van Galen et al., 2014). The importance of such regulation is emphasized by the observation that only a fraction of the variation in cellular protein levels can be explained through transcriptional changes (Jovanovic et al., 2015; Schwanhäusser et al., 2011). Importantly, studies of protein synthesis rates during hematopoiesis have indicated that dramatic changes occur during the early stages of lineage commitment (Signer et al., 2014). However, the functional consequences of such changes in protein synthesis rates for lineage commitment remain largely unexplored.

Diamond-Blackfan anemia (DBA) is a unique blood disorder where erythroid precursors and progenitors are selectively reduced in the bone marrow of patients, while all other lineages are ostensibly produced normally (Iskander et al., 2015; Nathan et al., 1978). Extensive studies have shown that the defect present in DBA appears to occur in early progenitors that are quantitatively reduced, but the few cells that do persist undergo normal terminal maturation (Nathan et al., 1978; Ohene-Abuakwa et al., 2005). The majority of DBA cases are caused by heterozygous loss-of-function mutations in one of 18 different ribosomal protein (RP) genes, resulting in RP haploinsufficiency (Mirabello et al., 2017). Despite extensive studies, the mechanisms by which a defect in RPs could cause a selective absence of erythroid cells within the hematopoietic compartment, while allowing for normal differentiation of other lineages, has remained a mystery (Sankaran and Weiss, 2015). Through studies of rare individuals with a diagnosis of DBA, we identified mutations in the key lineage-determining hematopoietic transcription factor GATA1 that can cause DBA (Sankaran et al., 2012). Motivated by these observations, we were able to show that RP haploinsufficiency results in reduced translation of *GATA1* mRNA and the erythroid defects present in DBA patient cells could largely be rescued by increasing GATA1 protein levels (Ludwig et al., 2014). However, despite this insight into the role of GATA1 in DBA pathogenesis, the mechanisms underlying such translational changes and the stages of hematopoiesis at which these alterations occur remain undefined.

DBA is a unique experiment of nature that presents an opportunity to better define the molecular mechanisms by which defects in the ribosome can selectively impact commitment to the erythroid, but not other hematopoietic lineages. Hence, mechanistic studies of DBA not only allow us to gain insight into the pathogenesis of this disease, but also provide us with an opportunity to better understand how protein translation may play a role in hematopoietic lineage commitment more generally. Here, we use human genetics to better define the role of ribosomal alterations *in vivo*, biochemical and proteomic studies to interrogate ribosome levels and composition in human hematopoietic cells, ribosome profiling in HSPCs undergoing erythroid lineage commitment to examine changes in global translation, deep transcriptome analysis of master regulators from unperturbed human HSPCs, and single-cell phenotypic analyses of primary DBA patient samples to define the mechanisms through which DBA arises and to gain insight into how translation plays a key role in the process of human hematopoietic lineage commitment. Importantly, we find that the quantity of ribosomes, but not the composition of such ribosomes, has a key role in promoting erythroid lineage commitment from HSPCs. Our work more generally reveals how ribosome levels can modulate cellular differentiation.

RESULTS

DBA Mutations in *TSR2* Highlight the Importance of Ribosome Production in Hematopoiesis

We reasoned that the identification of previously undefined genetic causes of DBA might provide additional insight into the underlying pathogenic mechanisms. By performing whole exome sequencing of DBA patients (Kim et al., 2017; Sankaran et al., 2012), we identified a hemizygous missense mutation in the X-linked and highly invariant *TSR2* gene in two male cousins with all the classical clinical features of DBA, as has been seen by others previously (Gripp et al., 2014) (Figures 1A and S1A; Tables S1 and S2). This finding piqued our interest, because the yeast ortholog of the RPS26 (eS26 in revised RP nomenclature) chaperone TSR2 has been shown to have an essential role in allowing productive formation of the mature ribosome and yet is biochemically distinct with complete nuclear localization (Schütz et al., 2014). Consistent with this, TSR2 was entirely localized to the nucleus in human hematopoietic cells (Figure S1B). Deletion of the yeast TSR2 ortholog results in a severe growth phenotype, which could be substantially rescued by introduction of human TSR2, but which had a reduced rescue by the allele observed in the two DBA patients (Figure 1B). This finding supports the contention that the *TSR2* mutation we identified results in a loss-of-function. Consistent with this, suppression of TSR2 through the use of short hairpin RNAs (shRNAs) was sufficient to impair erythroid lineage commitment and differentiation of human HSPCs (Figures 1C, 1D, and S1C). Furthermore, we observed phenotypes commonly seen with suppression of other genes implicated in DBA (Ludwig et al., 2014), including increased apoptosis, impaired growth, and a less mature erythroid gene expression profile, despite our use of cells with comparable global gene expression profiles (Figures S1D–S1G).

In agreement with our previous findings in DBA due to more typical RP gene mutations, TSR2 suppression resulted in selectively reduced levels of GATA1 protein, but did not affect the levels of *GATA1* mRNA (Figures 1E, 1F, and S1H–S1J). Increased expression of GATA1 protein in primary HSPCs with TSR2 suppression could rescue erythroid lineage commitment and differentiation (Figures 1G, 1H, S1K, and S1L). These data demonstrate that TSR2, which is biochemically unlinked from the mature ribosome and which has a key role in the production of adequate ribosome levels, is necessary for *in vivo* erythroid lineage commitment from human HSPCs. Considering these findings from a rare experiment of nature in addition to the more frequent RP mutations in DBA (Mirabello et al., 2017), we hypothesized that ribosome levels may have a selective role in human hematopoietic lineage commitment.

Molecular Lesions Underlying DBA Reduce Ribosome Levels in Hematopoietic Cells

Given the observations in yeast that the TSR2 ortholog is necessary for effective ribosome biogenesis and lesions in this gene reduce overall ribosome levels (Schütz et al., 2014), we wanted to interrogate the alterations in ribosome levels in HSPCs

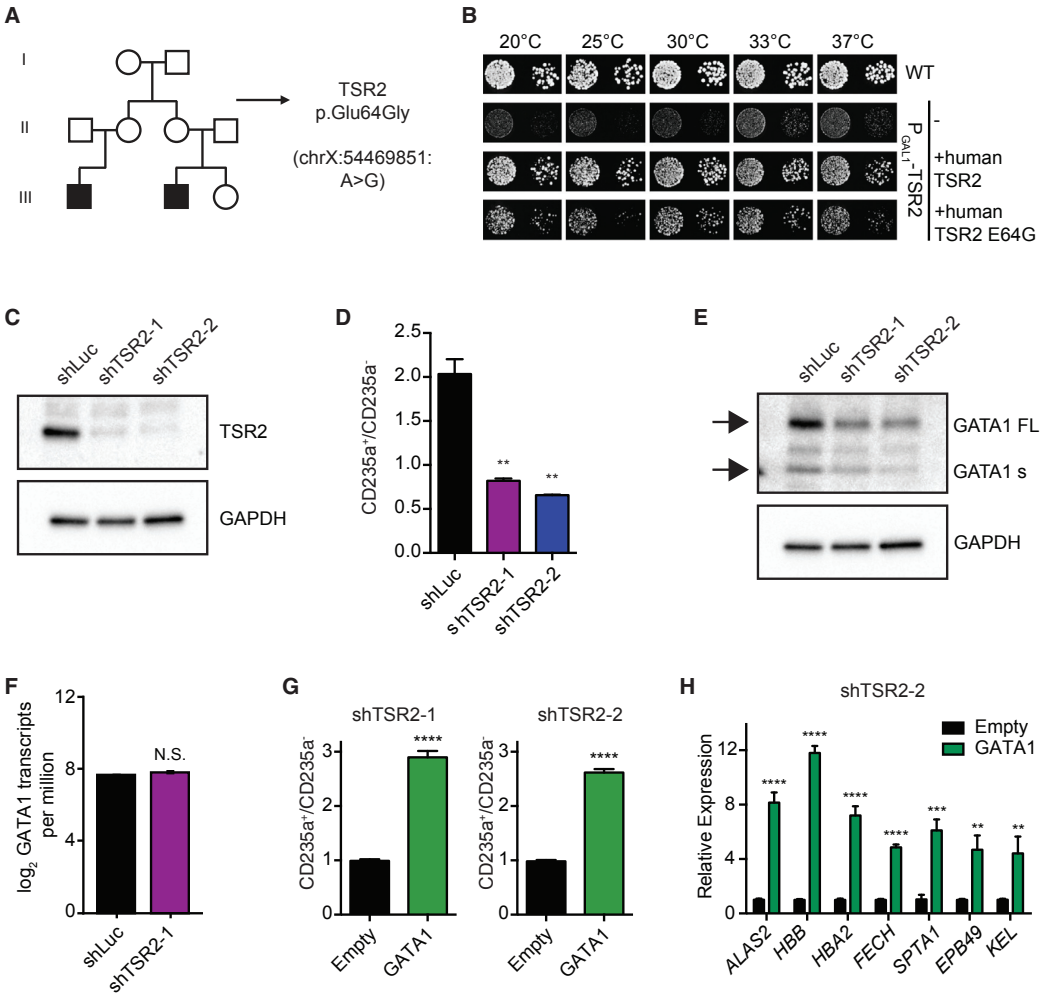


Figure 1. DBA with TSR2 Loss of Function

(A) Identification of a missense mutation in TSR2 in a pedigree with two affected male cousins.

(B) The human TSR2 ortholog could substantially rescue growth of the Tsr2-depleted yeast strain, while the TSR2 ortholog with the DBA-associated mutation had reduced rescue.

(C) Western blot showing the identification of two short hairpin RNAs (shRNAs) that target TSR2 in primary human HSPCs undergoing erythroid lineage commitment on day 5 after transduction.

(D) The ratio of erythroid ($CD235a^+$) to non-erythroid ($CD235a^-$) cells on day 5 in differentiating HSPCs after transduction with shRNAs targeting Luciferase (shLuc) or TSR2 (shTSR2). The data are shown as the mean ± SEM from three independent experiments. (**$p \leq 0.01$ using an unpaired two-tailed Student's t test).

(E) Western blot detection of GATA1 protein from lysates of differentiating HSPCs on day 5 after transduction. Arrowheads indicate GATA1 full length (FL) and GATA1 short (s), respectively, on top and bottom.

(F) *GATA1* mRNA levels derived from mRNA-seq in differentiating HSPCs. Shown is the mean ± SD of two biological replicates.

(G) The ratio of erythroid ($CD235a^+$) to non-erythroid ($CD235a^-$) cells on day 5 after transduction with shTSR2 and either a control vector or with GATA1 rescue. Shown is the mean ± the SD from three independent experiments. (****$p \leq 0.0001$ using an unpaired two-tailed Student's t test).

(H) Quantitative RT-PCR gene expression (normalized to β-actin) in differentiating HSPCs upon TSR2 suppression with or without GATA1 rescue. Shown is the mean ± the SD of three replicates. (**$p \leq 0.01$; ***$p \leq 0.001$; ****$p \leq 0.0001$ using an unpaired two-tailed Student's t test).

See also Figure S1 and Tables S1, S2, and S6.

undergoing commitment to the erythroid lineage and in hematopoietic cell lines. Similar to the characterized role of RPs in the biogenesis of mature ribosomes (Henras et al., 2015; Robledo et al., 2008), we found that suppression of TSR2 in human hematopoietic cells resulted in reduced levels of the 18S rRNA, with accumulation of its precursor, 18SE (Figures 2A, S2A, and S2B). Such a defect would impair production of the mature ribosome and thus limit the overall levels of ribosomes in the cytoplasm available for translation. Importantly, these defects are consistent with the lesions in ribosome maturation characterized

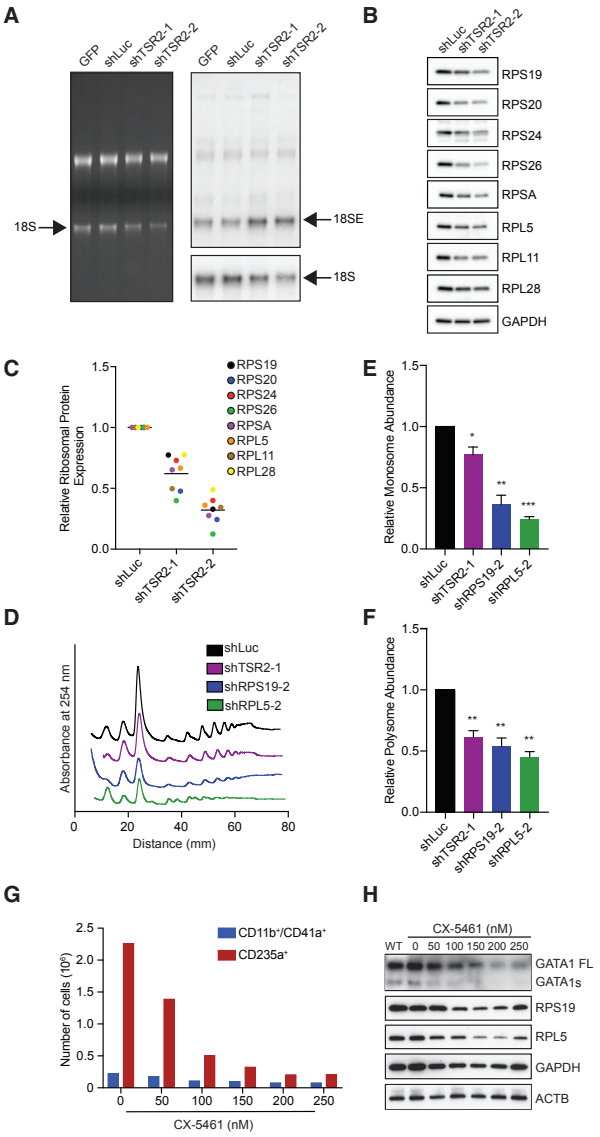


Figure 2. DBA-Associated Molecular Lesions Result in Reduced Ribosome Levels

(A) TSR2 suppression results in impaired pre-rRNA processing in human hematopoietic cells. Ethidium bromide-stained RNA gel (left) and Northern blot analysis (right) are shown in setting of TSR2 suppression.

(B) Western blot detection of the indicated proteins from lysates of differentiating HSPCs following TSR2 suppression.

(C) Relative quantification of RP intensities shown in (B) normalized to GAPDH.

(D) Polysome profiles of primary human HSPCs undergoing differentiation that show the reduction of monosome and polysome levels with DBA-associated molecular lesions. The traces are shown offset from one another on the arbitrary y axis (derived from relative absorbance at 254 nm) for ease of visualization.

(E and F) Relative quantification of monosome (E) and polysome (F) abundances from primary human HSPCs undergoing erythroid differentiation. Shown is the mean ± SD of two independent experiments. (*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$ using an unpaired two-tailed Student's t test).

in yeast with deletion of the TSR2 ortholog (Schütz et al., 2014). However, as the stages of ribosome biogenesis do vary between species (Preti et al., 2013), some difference in the precise nature of the impairment during this process is notable (Schütz et al., 2014).

We next investigated if there were alterations in the levels of actively translating ribosomes by creating DBA-associated molecular lesions, including suppression of TSR2, RPS19 (eS19), RPL5 (uL18), RPS24 (eS24), and RPL11 (uL5). We consistently observed reduced content of ribosomes in the cells using quantitative polysome profiling from similar numbers of cells and through quantification of a variety of RP levels in whole cell lysates from both primary hematopoietic cells and cell lines (Figures 2B–2D and S2C–S2Q). We found an overall reduction of 1.3- to 4.1-fold in the level of monosomes and 1.6- to 2.2-fold in the level of polysomes in primary hematopoietic cells (Figures 2E and 2F). This correlated well with the quantification of overall RP levels in these cells (Figures 2C and S2M–S2Q). Importantly, lesions in a single RP would generally suppress the protein levels of other RPs, particularly among those found in the same subunit as the primary molecular lesion (Figures 2B, 2C, and S2H–S2Q). These data collectively point toward an outcome of reduced ribosome levels with a diverse group of DBA-associated molecular lesions in differentiating HSPCs. To bolster these findings, given that rRNAs play a key role in the formation of the ribosome, we used a selective inhibitor of RNA polymerase I rRNA transcription (CX-5461) (Bywater et al., 2012) to show that rRNA inhibition more profoundly perturbed erythroid lineage commitment, as compared to other myeloid lineages and severely impaired GATA1 protein production (along with RPs) concomitantly (Figures 2G and 2H).

Verification of Constant Ribosome Composition in Human Hematopoietic Cells with DBA-Associated Molecular Lesions

Our results have suggested that molecular lesions resulting in DBA can reduce the level of actively translating ribosomes in human hematopoietic cells. These results in tandem with the *in vivo* findings from TSR2 mutant patients suggest, but do not formally prove, that reduced ribosome levels may be sufficient to result in impaired erythroid lineage commitment in HSPCs. Recent studies have suggested that RP mutations may result in altered ribosome composition in some contexts (Shi et al., 2017). We therefore wanted to directly interrogate the protein composition of actively translating ribosomes in the setting of DBA-associated lesions to understand whether such changes may occur in human hematopoietic cells. We performed quantitative high-coverage tandem-mass-tag (TMT) mass spectrometry in human

(G) Absolute numbers of erythroid cells as measured by surface marker expression of CD235a and myeloid cells as measured by CD41a or CD11b at 72 hr after treatment with increasing concentrations of the RNA polymerase I inhibitor CX-5461 in primary human HSPCs undergoing differentiation. Results from a representative experiment are shown.

(H) Western blot detection of the indicated proteins from lysates of differentiating HSPCs at 72 hr after treatment with increasing concentrations of CX-5461.

See also Figure S2.

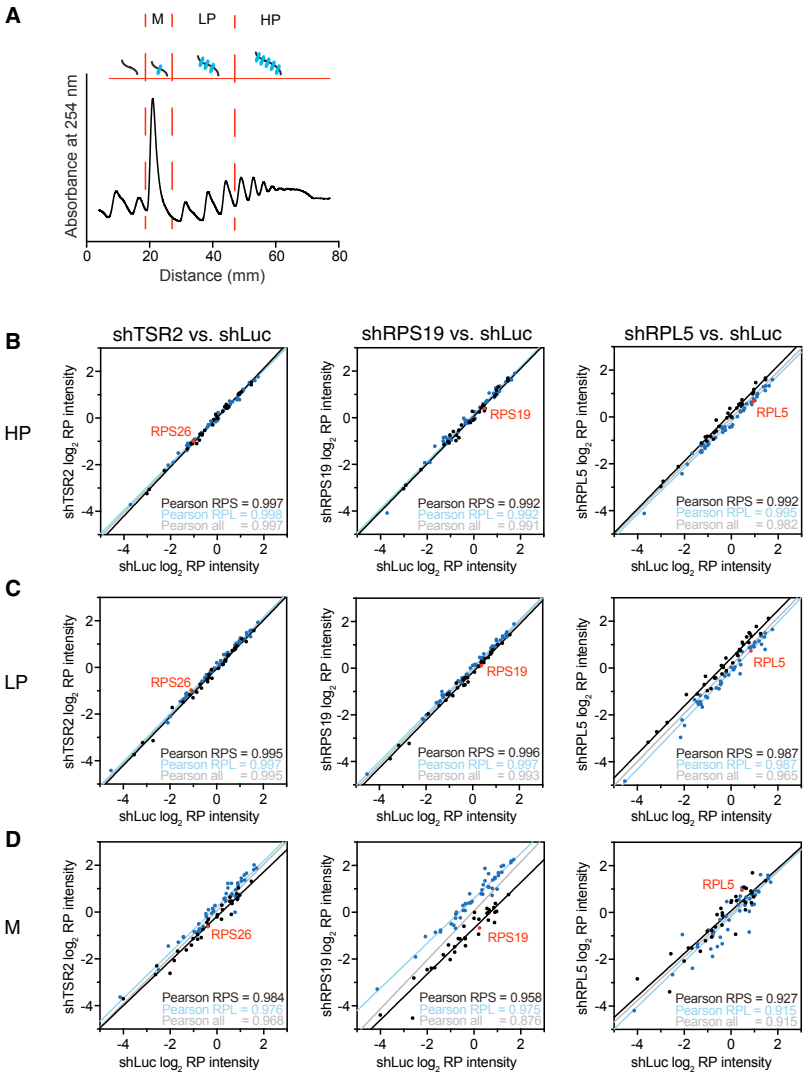


Figure 3. No Evidence for Variation in Ribosome Protein Composition in Cells with DBA-Associated Molecular Lesions

(A) Human hematopoietic cells treated with control vectors or with TSR2, RPS19, or RPL5 suppression were fractionated by sucrose gradient sedimentation. Monosome fractions (M), light polysomes (LP), and heavy polysomes (HP) were analyzed by tandem mass tag (TMT) mass spectrometry.

(B–D) Log_2 transformed and median centered RP intensities from two independent replicates in various knockdown (KD) conditions versus shLuc control in HP (B), LP (C), and M (D) fractions. RPs of the large subunit are shown in blue, RPs of the small subunit are shown in black, and the targeted or related RP is highlighted in red. Linear regressions for small subunit RPs (black), large subunit RPs (blue) and all RPs together (gray) are shown and Pearson correlations are reported.

See also Figure S3 and Table S7.

hematopoietic cells to measure the expression of all RPs. We fractionated cells by sucrose gradient sedimentation and collected monosomes (a single ribosome), light polysomes (2–4 ribosomes), and heavy polysomes (≥5 ribosomes) from control cells or those with DBA-associated perturbations, including haploinsufficiency of RPS19 and RPL5 or suppression of TSR2 (Figure 3A). Peptides for RPs were highly enriched in the mass spectrometry data: 77 out of 80 RPs were detectable by two or more unique peptides, and estimates of protein abundance were robust across biological replicates (Figures S3A–S3F). Strikingly, although we observed altered polysome profiles and cellular RP abundance, the average composition of RPs within monosomes, light polysomes, and heavy polysomes was largely invariant between controls and DBA-associated molecular lesions (Figures 3B–3D and S3G–S3I). The protein expression of the targeted or associated RPs did not deviate significantly from that of the other RPs (based on Studentized residuals, Figures S3G–S3I), strongly supporting the concept that DBA results from decreased ribosome abundance, rather than from formation of ribosomes that have a distinct protein composition. The composition of ribosome-associated proteins was also analyzed and we found no consistent alteration of these proteins in the presence of DBA-associated molecular lesions (Figures S3J and S3K). We note that because our assay measures total protein levels within a given cellular fraction, we cannot completely exclude the possibility that the pool of actively translating ribosomes is comprised of ribosomes with variable composition or that DBA-associated lesions could result in conformational changes in the ribosome that then alter translation. However, these possibilities seem unlikely, given the structural stability of the ribosome (Khatter et al., 2015) and the normal, albeit reduced, ribosomal maturation we observe. Therefore, our results from human genetic and biochemical studies of DBA-associated lesions lead to a model whereby the perturbation of hematopoietic differentiation observed arises from a reduced number of ribosomes per cell.

Defining Transcripts Whose Translation Is Most Sensitive to DBA-Associated Molecular Lesions

Having concluded that ribosome levels play a critical role in the lineage commitment defect observed in DBA from complementary human genetic and biochemical/proteomic studies, we aimed to better understand the consequences of decreased ribosome levels on translation. To gain global insight into changes in translation that occur with such perturbations in

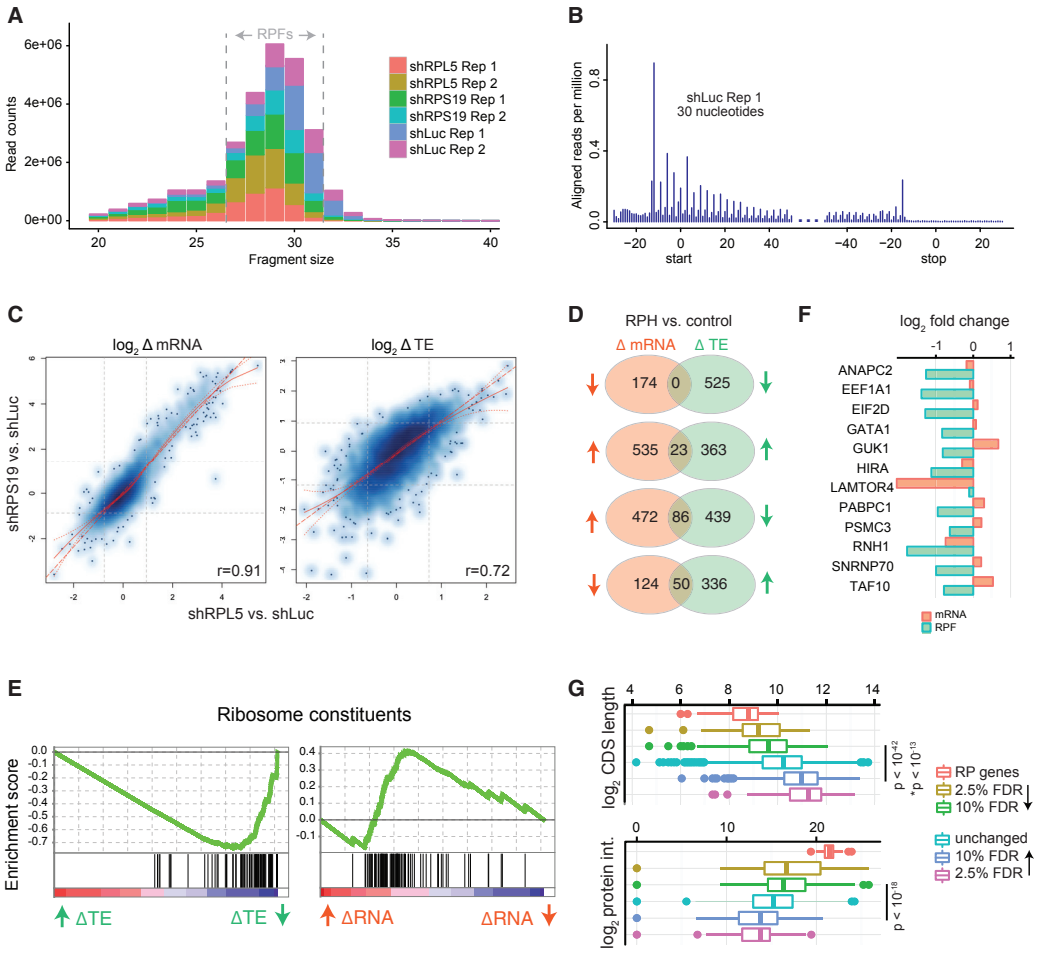


Figure 4. Identification of Transcripts Whose Translation Is Sensitive to RP Haploinsufficiency

(A) After adaptor trimming and rRNA removal, the distribution of ribosome profiling reads is shown. The reads all fall between 27–32 nucleotides.

(B) The ribosome profiling data exhibit triplet periodicity based upon meta-gene analysis of CDS regions. A representative example is shown.

(C) Differences between shLuc and shRPL5 or shRPS19 in primary differentiating human HSPCs are highly correlated at both the transcriptional and translational levels, as displayed in a scatterplot where color indicates point density. Both local regression (with confidence intervals) and linear fits are shown in red. Pearson correlations are indicated.

(D) Venn diagrams of differentially expressed (Δ mRNA, FDR <1% and $\log_2$ |fold change| >1) or differentially translated (Δ translation efficiency [TE], FDR <10%) genes showing that changes in translation and in transcription resulting from RP haploinsufficiency compared to control occur largely independent of each other.

(E) Gene set enrichment analyses indicate that RP genes are co-regulated at the translational (permutation FDR <0.0001), but not transcriptional (permutation FDR = 0.36) level with RP haploinsufficiency. The enrichment score is plotted in green, and genes are plotted as black lines according to their rank.

(F) The relative reduction in translation efficiency for selected RP haploinsufficiency-sensitive transcripts including GATA1 is shown in green, relative changes in mRNA expression are shown in red.

(G) Boxplots for CDS length or cellular protein intensities in primary human erythroid progenitors are shown across FDR thresholds for differential translation. CDS length was calculated for the most abundant transcript in shLuc and RP haploinsufficient differentiating HSPCs (*controlled for PolyA-selection based bias). p values were determined by an F-test.

See also Figure S4 and Tables S3 and S4.

primary human HSPCs undergoing erythroid lineage commitment, we performed ribosome profiling (Ingolia, 2016; Mills et al., 2016). This technique involves measuring translational efficiency (TE), by comparing the levels of ribosome-associated mRNA footprints to the total mRNA for each gene. For biological replicates of RPL5 and RPS19 suppression, we obtained both ribosome-protected footprints (RPFs) and matching mRNA-sequencing (mRNA-seq); the RPFs were of high quality, as assessed by expected RPF size, coding sequence (CDS) enrichment, and triplet periodicity (Figures 4A, 4B, S4A, and S4B; Table S3). Changes in transcription and translation appeared to be largely similar between RPS19 and RPL5 haploinsufficiency (Figure 4C), consistent with the concept that DBA-associated lesions cause a common set of molecular changes in

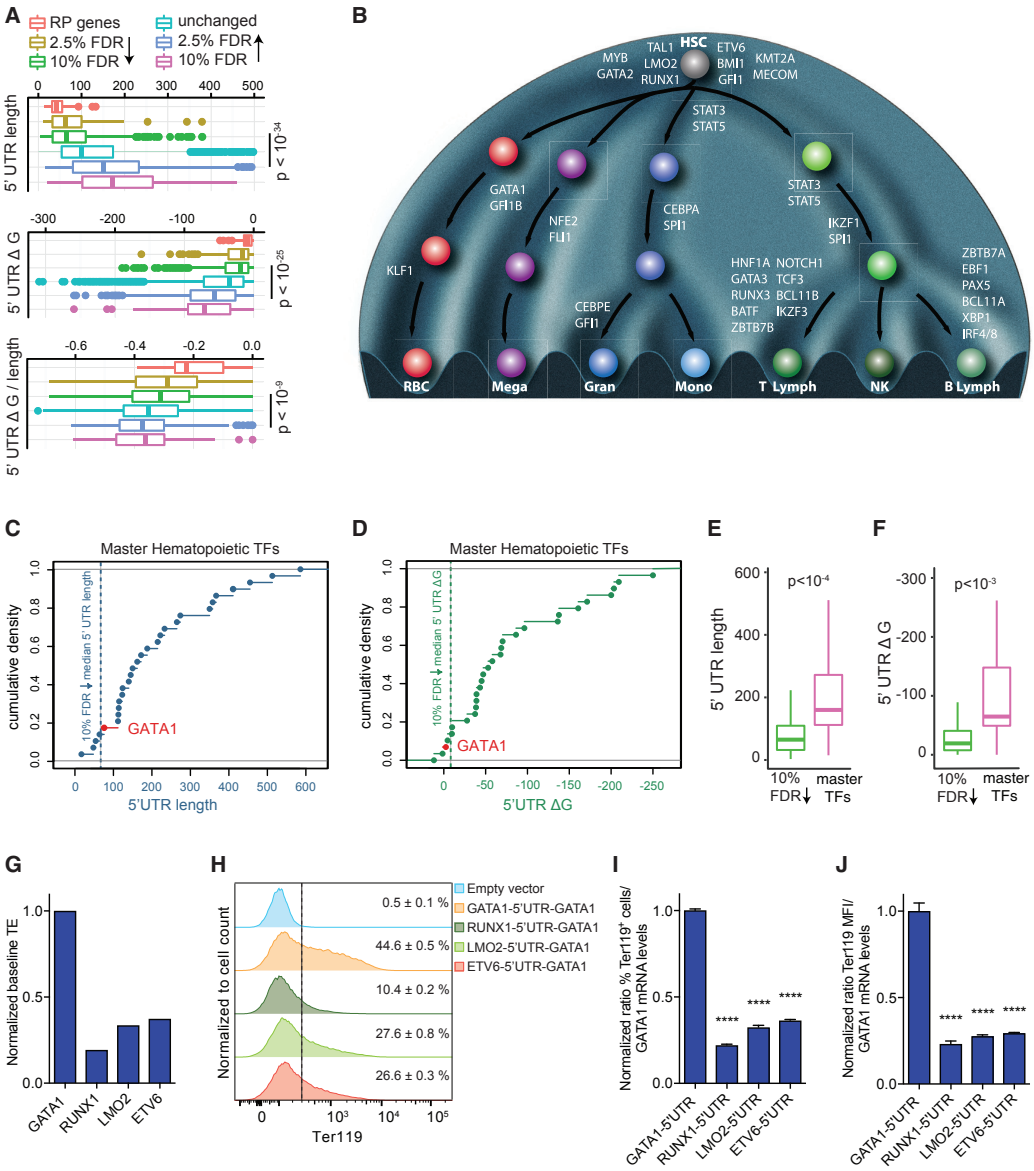


Figure 5. Analysis of 5′ UTR Features of Key Hematopoietic Transcription Factors

(A) Boxplots for different 5′ UTR features are shown across FDR thresholds for differential translation in primary differentiating human HSPCs. The minimum free energy (Δ G) was calculated using RNAfold for the entire 5′ UTR. As this prediction is correlated with length, ΔG corrected for 5′ UTR length was also analyzed. p values were determined by an F-test.

(B) Master regulator transcription factors (TFs) are shown in their approximate positions of action in a model of hematopoiesis. HSC, hematopoietic stem cell; RBCs, red blood cells; Mega, megakaryocyte; Gran, granulocyte; Mono, monocyte; B Lymph, B lymphocyte; T Lymph, T lymphocyte; NK, natural killer cell.

(C and D) The GATA1 5′ UTR is shorter (C) and less structured (D) than those of most other hematopoietic master TFs. GATA1 is highlighted in red. The median line for the 10% FDR RP haploinsufficiency-sensitive transcripts is indicated, respectively.

(E and F) Most hematopoietic master TFs have significantly longer (2.5 mean-fold difference) (E) and more structured 5′ UTRs (2.8 mean-fold difference in ΔG) (F) than transcripts that are translationally downregulated with RP haploinsufficiency.

(G) Normalized baseline translation efficiencies (TE) based on ribosome profiling in unperturbed HSPCs undergoing erythroid lineage commitment are shown for GATA1, RUNX1, LMO2, and ETV6.

(H) Histogram plots for Ter119 in GFP$^+$ populations derived from G1E cells that were transduced with GATA1-, RUNX1-, LMO2-, or ETV6-5′UTR-GATA1 cDNA constructs. The mean ± the SD for the percentages of Ter119$^+$ cells of three replicates is shown.

(legend continued on next page)

human HSPCs undergoing erythroid differentiation. Importantly, changes in transcription and translation were largely independent (Figures 4D and S4C), emphasizing the value of ribosome profiling (Ingolia, 2016).

Notably, the RP genes globally showed the greatest decrease in TE with RP haploinsufficiency (the top 10 of 557 KEGG, REACTOME, and BIOCARTA pathways are primarily composed of RP genes), despite relatively unchanged mRNA levels (Figure 4E). This observation suggests that RPs are co-regulated at the translational level, which would allow cells to maintain RP stoichiometry. While translational co-regulation of RPs has been demonstrated downstream of mTOR signaling (Hsieh et al., 2012; Thoreen et al., 2012), our findings show, similar to observations made in yeast (Thompson et al., 2016), that co-regulation of RP translation can also occur in the setting of RP haploinsufficiency in human hematopoietic cells. The extent to which the ubiquitin-dependent degradation of RPs (Sung et al., 2016) plays an additional role in maintaining homeostasis is unclear. However, our data suggest that the reduction of ribosome levels observed in the setting of RP haploinsufficiency (Figures 2 and S2) is largely promoted through reduced translation of RP mRNAs.

At a false discovery rate (FDR) of 10%, we identified a selective set of 525 transcripts whose TE was particularly sensitive to and downregulated by RP haploinsufficiency (Figure 4D; Table S4). We confirmed our previous finding that translation of *GATA1* mRNA is significantly decreased by ~2-fold in differentiating HSPCs with RP haploinsufficiency (Ludwig et al., 2014) (Figure 4F; Table S4). A subset of the downregulated transcripts are essential for growth in hematopoietic cells (Wang et al., 2015) and are substantially upregulated during early erythropoiesis (between $CD34^+$ and proerythroblast (ProE) stages of normal human erythropoiesis) (Li et al., 2014), consistent with the stages of perturbation observed in DBA patients (Figure S4D; Table S4). This observation suggests that the reduced translation of multiple transcripts that are upregulated at the early stages of erythroid lineage-specification from HSPCs, including *GATA1*, plays a key role in the *in vivo* phenotypes observed in DBA. Importantly, in this context, we note that mutations in *GATA1* are sufficient to cause DBA in rare patients (Sankaran et al., 2012) and some genes that are downregulated at the translational level, such as the ribosome-associated protein RNH1, have been shown to have additional key roles in the regulation of *GATA1* mRNA translation (Chennupati et al., 2018). In concert with previous genetic and rescue experiments performed in DBA patient samples (Ludwig et al., 2014), our results suggest that a number of ribosome-associated factors are translationally downregulated in the setting of RP haploinsufficiency and many of these lesions potentially result in the coordinated impairment of *GATA1* mRNA translation as a common downstream pathogenic mechanism.

We next sought to determine if these RP haploinsufficiency-sensitive transcripts shared similar features to gain insight into the mechanisms of lineage commitment during human hematopoiesis and how this process can be perturbed in diseases like DBA. Interestingly, we found that the RP haploinsufficiency-sensitive transcripts were on average shorter in overall length, more efficiently translated under baseline conditions, and encoded more abundantly expressed proteins in unperturbed primary human erythroid progenitors (Gautier et al., 2016) (Figures 4G and S4D). Of note, short mRNA length has been shown to be associated with efficient translation in other contexts (Thompson et al., 2016), although this feature alone may not be sufficient to mediate translational control.

Much of the underlying regulation of protein translation is mediated by the 5′ untranslated region (5′ UTR) of transcripts (Hinnebusch et al., 2016; Shah et al., 2013). To fully interrogate this variation, we comprehensively defined 5′ UTRs present in hematopoietic cells using cap analysis gene expression (CAGE) sequencing, which can often vary from annotated 5′ UTRs (Figure S5A). Using such data, we found that the 5′ UTRs of downregulated transcripts were 42 nucleotides shorter on average, were predicted to have less complex secondary structure, and contained fewer in-frame and out-of-frame upstream start codons (uAUGs)—features associated with efficient ribosome initiation and translation in unperturbed cells, including in our data from control HSPCs undergoing erythroid lineage commitment (Hinnebusch et al., 2016; Shah et al., 2013) (Figures 5A, S5B, and S5C). As the 5′ terminal oligopyrimidine (5′ TOP) motif was originally identified in RP mRNAs (Roepcke et al., 2006), we investigated whether this motif or a similar motif was enriched in those transcripts with reduced TE. We found a significant enrichment for such motifs that was predominantly explained by the downregulated group of RP mRNAs (Figure S5D), suggesting that translational alterations in RP haploinsufficiency are partially overlapping with, but are distinct from, alterations due to mTOR inhibition where TOP or TOP-like motifs are present in a large subset of mTOR-sensitive transcripts (Hsieh et al., 2012; Thoreen et al., 2012). Further analysis revealed that a number of motifs were nominally enriched across the entire 5′ UTR, as well as at the 5′ and 3′ ends, but no single motif could explain the observed differences in TE between RP haploinsufficiency-sensitive and insensitive transcripts (Figures S5E and S5F). Altogether, a model of the features investigated here explained 39% of the variation of TE changes in a held-out set of genes, validating the key role that these features have in translational regulation (Hinnebusch et al., 2016; Shah et al., 2013).

Interrogation of 5′ UTRs from Hematopoietic Master Regulators Suggests Mechanisms of Lineage Selectivity in DBA

While our ribosome profiling analysis elucidated transcripts within differentiating HSPCs that selectively show increased sensitivity to impaired translation in the setting of DBA-associated molecular lesions, these findings are insufficient to explain

(I) Bar graphs for normalized ratios of % $Ter119^+$ populations in GFP^+ cells/*GATA1* mRNA levels from G1E cells that were transduced with GATA1-, RUNX1-, LMO2-, or ETV6-5′UTR-GATA1 constructs. The mean ± the SD of three replicates is shown (****$p \leq 0.0001$ using an unpaired two-tailed Student's t test).
(J) Bar graphs for normalized ratios of the Ter119 mean fluorescence intensities (MFIs) of GFP^+ cells/*GATA1* mRNA levels from G1E cells that were transduced with the constructs listed above. The mean ± the SD of three replicates is shown (****$p \leq 0.0001$ using an unpaired two-tailed Student's t test).
See also Figure S5 and Table S5.

the erythroid specificity of DBA. We had noted that most of the transcripts sensitive to RP haploinsufficiency tended to have short and unstructured 5′ UTRs—features that are associated with increased translation efficiency under baseline conditions (Figures 5A and S5B). This included the 5′ UTR of *GATA1* mRNA. Master regulator transcription factors, such as GATA1, are critical for determining cell identity and promoting lineage specification in physiologic differentiation processes such as hematopoiesis (Doulatov et al., 2012; Orkin and Zon, 2008). Indeed, such master regulator transcription factors are sufficient to allow for dramatic changes in cell state (Srivastava and DeWitt, 2016; Capellera-Garcia et al., 2016). We reasoned that perhaps the observed lineage selectivity may occur because no other master regulators of hematopoietic lineage commitment were perturbed by reduced ribosome levels (Paul et al., 2015; Velten et al., 2017). Our results from ribosome profiling suggest that this could be due to 5′ UTR-mediated mechanisms. To investigate whether the observed patterns of sensitivity to reduced ribosome levels may underlie the hematopoietic lineage selectivity, we examined CAGE data generated from unperturbed primary human HSPCs that are comprised of progenitors capable of commitment to multiple lineages. Among a group of 36 well-characterized hematopoietic master regulators known to have key and well-defined roles in lineage commitment (where 29 of these transcription factors were well expressed and had clearly defined TSSs in CAGE data generated from unperturbed primary human HSPCs) (Figure 5B), we found that the majority had significantly longer and more complex 5′ UTRs compared with those transcripts sensitive to reduced ribosome levels, with *GATA1* mRNA being a notable exception (Figures 5C and 5D). Importantly, the overall group of hematopoietic master regulators has significantly longer 5′ UTR lengths (2.5 mean-fold difference, $p < 10^{-4}$) and more complex 5′ UTR structures (2.8 mean-fold difference in ΔG, $p < 10^{-3}$) than the group of transcripts showing sensitivity to RP haploinsufficiency (Figures 5E and 5F). Altogether, these data suggest that *GATA1* exhibits unique 5′ UTR features among hematopoietic master regulators, which may explain its translational sensitivity to reduced ribosome levels and the consequent lineage-specific defect observed in DBA. Importantly, we were able to validate this lack of translational downregulation with RP haploinsufficiency for master regulators that were expressed in the differentiating HSPCs: *KLF1*, *TAL1*, *MYB*, *GATA2*, *LMO2*, *RUNX2*, *ETV6*, *KMT2A*, *NFE2*, *FLI1*, *STAT5A*, *STAT3*, *SPI1*, *NOTCH1*, *BCL11A*, *IKZF1*, and *XBP1* all showed no major decrease in TE (FDR↓ >10%, $\log_2$ TE fold decrease of <0.45).

To directly interrogate whether such 5′ UTR features may be sufficient to confer baseline variation in translation, we complemented the GATA1 null G1E hematopoietic cell line (Weiss et al., 1997) with *GATA1* cDNA harboring 5′ UTRs from different hematopoietic master regulators including *GATA1* itself, *LMO2*, *RUNX1*, and *ETV6*—the latter three being longer and having more complex secondary structures than the endogenous *GATA1* 5′ UTR (Table S5). Consistent with the hypothesis that other master regulator 5′ UTRs should have lower translation efficiency under baseline conditions (Figure 5G) and therefore would be less susceptible to a reduction in ribosome levels, we found that the GATA1-induced erythroid differentiation (that correlates with GATA1 protein levels) was substantially impaired by 5′ UTRs from the other hematopoietic master regulators compared with *GATA1* (Figures 5H–5J). These data emphasize the unique features of the *GATA1* 5′ UTR, in comparison to other hematopoietic master regulator mRNAs, which thereby confer sensitivity to variation in ribosome levels.

Impaired GATA1 Protein Production in Primary HSPCs from DBA Patients

We have shown that impaired translation of select transcripts, including *GATA1*, occurs with RP haploinsufficiency and consequently reduced ribosome levels and is accompanied by the functional hematopoietic defects characteristic of DBA. Our analysis suggests that a key common effector of these defects in DBA is GATA1. We wanted to confirm the relevance of these findings at the single cell level in hematopoietic progenitors *in vivo* in DBA patients. As primary patient samples are often limited and challenging to obtain, we developed a semiquantitative immunohistochemistry staining method for GATA1 protein expression (Lee et al., 2017). We could individually identify and measure the staining intensity of GATA1 in the nuclei of erythroid precursors and progenitors from bone marrow biopsies obtained from healthy controls or from DBA patients (Carpenter et al., 2006). We found that DBA patients had a significantly reduced GATA1 staining intensity in such cells (Figures 6 and S6). While some cells did have overlapping intensities, we noted that less mature cells with larger nuclei frequently had reduced staining intensities, suggesting that the defects in DBA arise at the early stages of erythroid lineage commitment. However, immunohistochemistry is limited in our ability to compare stage-matched cells and this analysis could be confounded by variation in erythroid cell composition between DBA patients and controls.

We therefore wanted to identify the stages at which such impairments may arise during *in vivo* human hematopoiesis. Recent work has shown that lineage commitment to the erythroid and other lineages occurs predominantly at the early HSPC stages, rather than occurring at later stages of differentiation as classically inferred through analysis of heterogeneous bulk cell populations (Notta et al., 2016; Paul et al., 2015; Perié et al., 2015; Velten et al., 2017). Indeed, *GATA1* mRNA shows initial expression in human HSPCs within the most primitive $CD34^+CD38^-$ compartment (Notta et al., 2016). We had previously demonstrated that human HSPCs show no difference in *GATA1* mRNA expression when comparing healthy donors to patients with DBA (Ludwig et al., 2014). To interrogate GATA1 protein expression at the single cell level, we developed an intracellular flow cytometric detection approach. We utilized an *in vitro* erythroid differentiation protocol from human HSPCs to interrogate GATA1 expression during this differentiation process (Giani et al., 2016; Kim et al., 2017). We found that GATA1 was expressed at low levels in a subset of the HSPCs prior to initiation of differentiation. As expansion and differentiation proceeded, there was an initial upregulation of GATA1 in many cells and a progressive increase in expression among the primitive $CD34^+CD38^-$ and more differentiated $CD34^+CD38^+$ HSPC populations (Figure 7A). With differentiation, robust and high-level GATA1 protein expression was seen in lineage committed $CD235a^+CD71^+$ erythroid cells (Figures 7A and S7A). Our

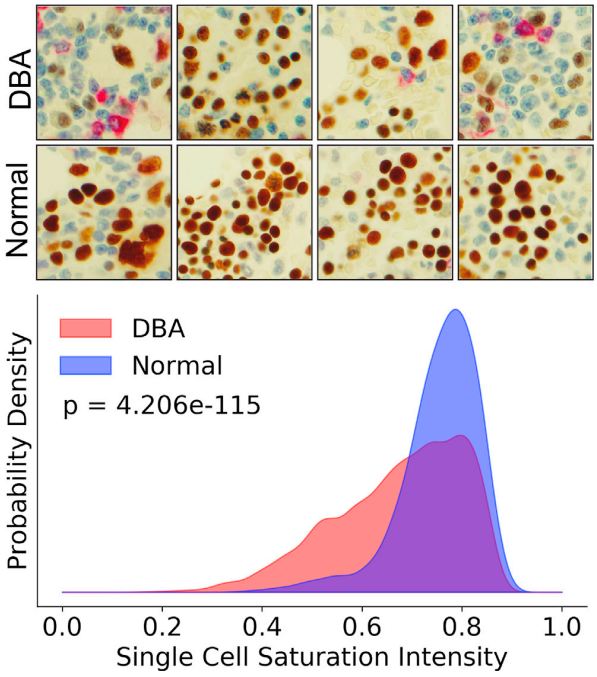


Figure 6. Reduced GATA1 Protein Levels in Bone Marrow Progenitors from DBA Patients

Representative images of human bone marrow biopsies stained for GATA1 protein (brown) in DBA patients with diverse RP mutations and normal healthy controls. Below, is a density plot comparing single cell saturation intensities between DBA patients and normal individuals that shows significantly reduced expression in DBA (n = 2,759 for DBA and 2,149 cells for controls; significance calculated by the Mann-Whitney U test).
See also Figure S6.

findings from this differentiation protocol demonstrate that GATA1 is initially expressed in a subset of HSPCs at low levels and this expression then progressively increases with higher-level expression occurring in erythroid-committed progenitors and precursors.

The observed early expression of GATA1 protein in HSPCs is consistent with the recently described models of hematopoiesis where lineage commitment occurs in such primitive populations and builds upon these findings to delineate a key role for this master transcription factor in this process (Notta et al., 2016; Paul et al., 2015; Perié et al., 2015; Velten et al., 2017). Given the observed expression, we interrogated GATA1 protein expression in single cells from HSPC populations of unperturbed DBA patient or healthy control bone marrow aspirate samples (Figures 7B and 7C). Interestingly, among patients with *RPL35A*, *RPL5*, or *RPS19* mutations, there was a consistent reduction in GATA1 expression in both $CD34^+CD38^-$ and $CD34^+CD38^+$ HSPC populations (Figure 7C). Despite overall upregulation of GATA1 protein levels during the $CD38^-$ to $CD38^+$ transition of hematopoietic progenitors, the overall GATA1 levels in individual progenitors remained lower in DBA patients. These observations demonstrate that in uncultured bone marrow specimens from DBA patients with diverse RP mutations, there is a reduction in GATA1 expression at the early HSPC stages. This finding fits with the lineage commitment impairment characteristic of DBA (Iskander et al., 2015; Nathan et al., 1978) and also supports our mechanistic studies of altered translation in differentiating HSPCs.

DISCUSSION

Recent studies have refined our understanding of hematopoiesis and shown that hematopoietic lineage commitment occurs at the early HSPC stages (Notta et al., 2016; Paul et al., 2015; Perié et al., 2015; Velten et al., 2017). However, the key molecular regulators of lineage commitment and the relevance of these updated models to human disease have not been explored. Here, we have studied a rare genetic blood disorder—DBA—that is characterized by a paucity of erythroid precursors and progenitors, to provide insight into both of these issues. We show that the lesions in DBA arise at the level of HSPCs, consistent with the specification of lineage commitment and differentiation within this primitive cell compartment. We also demonstrate how ribosome levels can play a key role in allowing lineage commitment to productively ensue. Our findings demonstrate how by exploring a rare genetic disorder, we can not only gain insight into the pathogenesis of the specific disease of interest, but also more broadly provide insight into the molecular underpinnings of hematopoietic lineage commitment.

We demonstrate through complementary human genetic and biochemical studies that ribosome levels serve a key role in allowing effective hematopoietic differentiation. A select subset of transcripts is affected by functionally relevant alterations in ribosome levels. Specifically, we found that reduced ribosome levels impaired the translation of transcripts that are normally highly translated and have short/unstructured 5′ UTRs over other transcripts. These findings demonstrate the value that ribosome profiling can have to interrogate translation on a global genomic scale and have allowed us to identify the specific liabilities that occur in the setting of reduced ribosome levels (Ingolia, 2016). Our findings complement recent studies showing how protein synthesis undergoes dramatic variation during hematopoiesis (Signer et al., 2014). While the functional role of such tightly regulated protein synthesis rates in hematopoietic stem cells has been examined, the necessity of upregulation in protein synthesis rates for hematopoietic differentiation has not been explored. While in some contexts RP composition may vary (Shi et al., 2017), we find that in the setting of RP haploinsufficiency in hematopoietic cells, no apparent altered composition can be identified. Rather, the impaired lineage commitment characteristic of DBA arises from a reduced cellular level of ribosomes. It is notable that studies in hematopoietic cells have demonstrated that the highest rates of protein synthesis occur in progenitors undergoing erythroid lineage commitment (Signer et al., 2014), which fits with our findings of how ribosome levels can selectively impair erythroid lineage commitment and GATA1 requires one of the highest translation rates among various master regulators of hematopoiesis. Future studies examining the sensitivity to and liabilities arising from reduced ribosome levels in various hematopoietic lineages will provide further insight into this process.

Beyond hematopoiesis, the regulation of ribosome levels is likely to have a key role more broadly in cellular differentiation

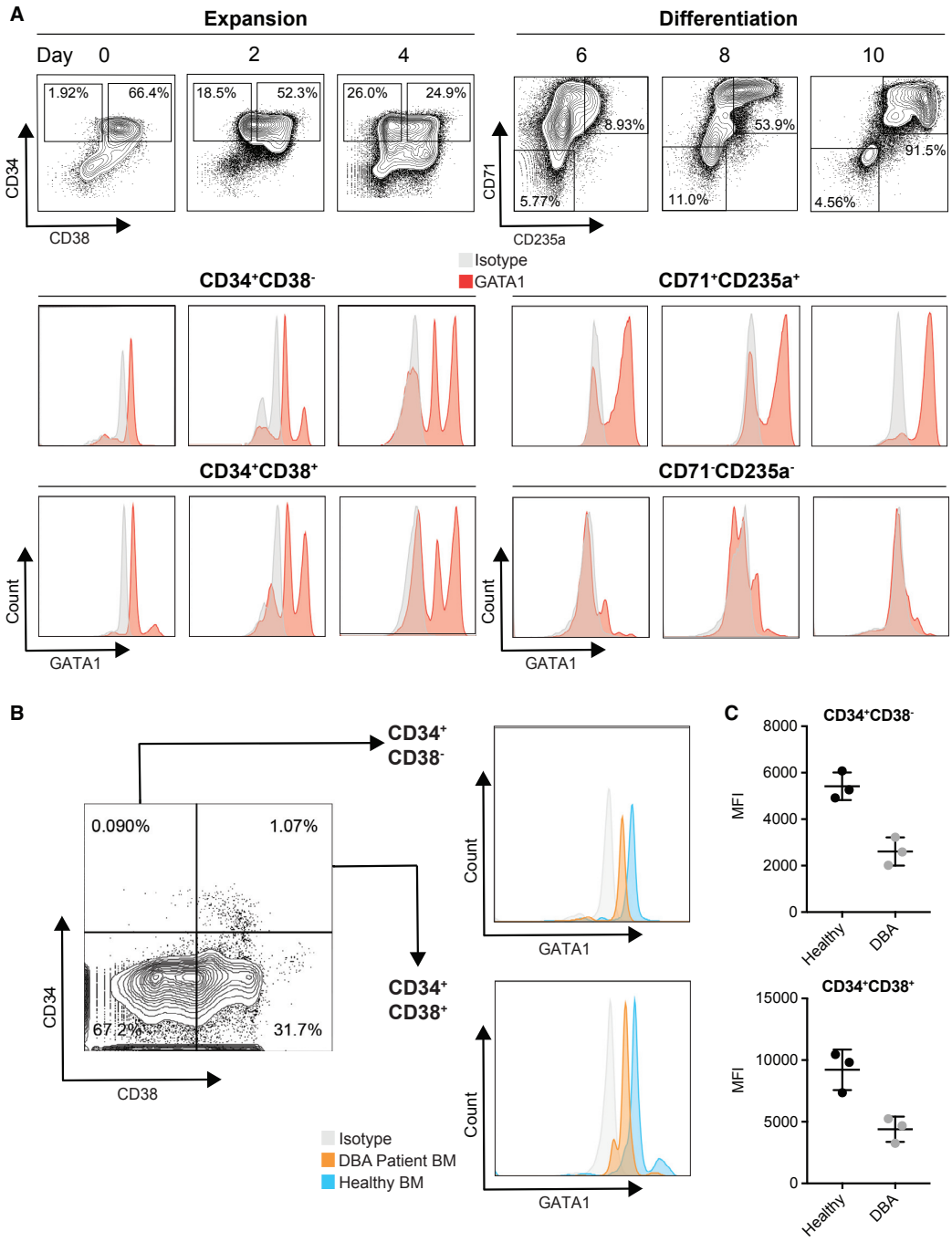


Figure 7. Reduced GATA1 Protein Expression in Primary HSPCs from DBA Patients

(A) Intracellular flow cytometric detection shows low levels of GATA1 expression in a subset of both the primitive $CD34^{+}CD38^{-}$ and more differentiated $CD34^{+}CD38^{+}$ HSPC populations (left). With differentiation, robust and high-level GATA1 protein expression can be seen in committed $CD235a^{+}CD71^{+}$ erythroid cells (right).

(B) Reduced GATA1 protein expression in single cells from HSPC populations from a DBA patient bone marrow aspirate sample compared to a healthy control.

(C) GATA1 MFIs show a consistent reduction in GATA1 expression in $CD34^{+}CD38^{-}$ and $CD34^{+}CD38^{+}$ HSPC populations in DBA patients with *RPL35A*, *RPL5*, or *RPS19* mutations compared to healthy controls.

See also Figure S7.

and tissue homeostasis (Buszczak et al., 2014). Indeed, the broad array of ribosomal disorders, which display highly-specific phenotypes, indicates the key roles that ribosome levels may have in other contexts and cell types (McCann and Baserga, 2013). Even mutations in RPs themselves can present with a broad range of highly specific phenotypes beyond the paucity of erythroid cells characteristic of DBA. These phenotypes include isolated congenital asplenia (Bolze et al., 2013) and neurodevelopmental disorders (Brooks et al., 2014) in addition to the non-hematopoietic phenotypes notable in DBA patients, such as cleft lip/ palate, thumb abnormalities, and other congenital defects (Gazda et al., 2008). It is likely that mechanisms involving impaired translation of specific transcripts, similar to those we identify within the hematopoietic compartment, may have a role in mediating these other phenotypes.

While studies of cellular differentiation have largely focused on transcriptional changes underlying these processes, it is clear that post-transcriptional regulation serves key and largely unappreciated roles in this process. While exploration of such mechanisms is more limited, as compared to the relative ease of interrogating the transcriptome (Tanay and Regev, 2017), advances in approaches such as ribosome profiling suggest that important insight can more broadly be gained into this process through in depth mechanistic studies (Ingolia, 2016). With continued advances in the ability to carry out such approaches in more limited populations of cells, as we have done here with primary human hematopoietic cells, and the increased availability of orthogonal genomic data, more sophisticated insight can be gained into the regulation of this process. In addition, the key advances occurring in the field of human genetics will enable us to better understand how such process can be perturbed in human disease (Casanova et al., 2014), as we have been able to study here for DBA.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Primary Cell Culture
  - 293T and K562 Cell Culture
  - G1E Cell Culture
- METHOD DETAILS
  - Lentiviral Vectors and Infection
  - Yeast Strains and Plasmids
  - Quantitative RT-PCR
  - Western Blotting
  - Flow Cytometry Analysis and Apoptosis Detection
  - Intracellular GATA1 Staining
  - rRNA Processing Examination
  - Polysome Profiling
  - Mass Spectrometry
  - RNA Polymerase I Inhibition
  - Ribosome Profiling
  - 5′UTR-GATA1 Construct Cloning
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Whole Exome Sequencing
  - ExAC Gene Constraint Analyses
  - Analysis of Mass Spectrometry Data
  - Analysis of RNA and Ribosome Profiling Libraries
  - Re-Annotation of 5′ UTRs
  - Analysis of Features for Association with ΔTE
  - Motif Analyses
  - Bone Marrow Biopsy Section Immunohistochemical Staining and Analysis
  - Statistical Analyses
- DATA AND SOFTWARE AVAILABILITY
  - Accession Codes

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and seven tables and can be found with this article online at https://doi.org/10.1016/j.cell.2018.02.036.

ACKNOWLEDGMENTS

We thank D. Nathan, S. Orkin, L. Zon, K. Patel, S. Eichhorn, Z. Ji, J. Clohessy, A. Bolze, and Sankaran laboratory members for valuable discussions. We are grateful to T. DiCesare for assistance with illustrations. R.K.K received partial support from a Boehringer Ingelheim MD Fellowship. S.K.M. received support from the NIH (T32 HL007574). A.R. is an investigator of the Howard Hughes Medical Institute. V.G.P. received support from the Swiss National Science Foundation, Novartis Foundation, Olga Mayenfisch Stiftung, and the European Research Council (EURIBIO260676). V.G.S. is a Principal Faculty member of the Harvard Stem Cell Institute. This work was supported by the NIH (R01 DK103794 and R33 HL120791), as well as a grant from the DBA Foundation and a March of Dimes Basil O'Connor Scholar Award (to V.G.S.).

AUTHOR CONTRIBUTIONS

Conceptualization, R.K.K., J.C.U., and V.G.S.; Methodology, R.K.K., M.M., J.C.U., C.F., L.S.L., S.K.M., and V.G.S.; Formal Analysis, R.K.K., M.M., J.C.U., C.F., L.S.L., S.K.M., D.R.M., and V.G.S.; Investigation, R.K.K., M.M., J.C.U., C.F., L.S.L., S.K.M., N.J.A., H.S., H.K., D.R.M., M.J., S.R.E., C.P.F., J.M.E., S.S., J.L., K.W.G., O.K.W., G.S.P., L.G., A.R., E.S.L., H.T.G., W.Y.L., V.G.P., S.A.C., and V.G.S.; Resources, J.C.U., S.K.M., J.M.E., and V.G.S.; Writing – Original Draft, R.K.K., M.M., J.C.U., and V.G.S.; Writing – Review & Editing, R.K.K. and V.G.S. with input from all authors; Visualization, R.K.K., M.M., J.C.U., C.F., L.S.L., S.K.M., and V.G.S.; Supervision, A.R., E.S.L., S.A.C., and V.G.S.; Funding Acquisition, V.G.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 24, 2017
Revised: November 28, 2017
Accepted: February 15, 2018
Published: March 15, 2018

REFERENCES

An, X., Schulz, V.P., Li, J., Wu, K., Liu, J., Xue, F., Hu, J., Mohandas, N., and Gallagher, P.G. (2014). Global transcriptome analyses of human and murine terminal erythroid differentiation. Blood *123*, 3466–3477.

Arner, E., Daub, C.O., Vitting-Seerup, K., Andersson, R., Lilje, B., Drabløs, F., Lennartsson, A., Rönnerblad, M., Hrydziuszko, O., Vitezic, M., et al.; FANTOM Consortium (2015). Transcribed enhancers lead waves of

coordinated transcription in transitioning mammalian cells. Science *347*, 1010–1014.

Bolze, A., Mahlaoui, N., Byun, M., Turner, B., Trede, N., Ellis, S.R., Abhyankar, A., Itan, Y., Patin, E., Brebner, S., et al. (2013). Ribosomal protein SA haploinsufficiency in humans with isolated congenital asplenia. Science *340*, 976–978.

Brooks, S.S., Wall, A.L., Golzio, C., Reid, D.W., Kondyles, A., Willer, J.R., Botti, C., Nicchitta, C.V., Katsanis, N., and Davis, E.E. (2014). A novel ribosomopathy caused by dysfunction of RPL10 disrupts neurodevelopment and causes X-linked microcephaly in humans. Genetics *198*, 723–733.

Buszczak, M., Signer, R.A., and Morrison, S.J. (2014). Cellular differences in protein synthesis regulate tissue homeostasis. Cell *159*, 242–251.

Bywater, M.J., Poortinga, G., Sanij, E., Hein, N., Peck, A., Cullinane, C., Wall, M., Cluse, L., Drygin, D., Anderes, K., et al. (2012). Inhibition of RNA polymerase I as a therapeutic strategy to promote cancer-specific activation of p53. Cancer Cell *22*, 51–65.

Capellera-Garcia, S., Pulecio, J., Dhulipala, K., Siva, K., Rayon-Estrada, V., Singbrant, S., Sommarin, M.N., Walkley, C.R., Soneji, S., Karlsson, G., et al. (2016). Defining the minimal factors required for erythropoiesis through direct lineage conversion. Cell Rep. *15*, 2550–2562.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. *7*, R100.

Casanova, J.L., Conley, M.E., Seligman, S.J., Abel, L., and Notarangelo, L.D. (2014). Guidelines for genetic studies in single patients: lessons from primary immunodeficiencies. J. Exp. Med. *211*, 2137–2149.

Chennupati, V., Veiga, D.F., Maslowski, K.M., Andina, N., Tardivel, A., Yu, E.C., Stilinovic, M., Simillion, C., Duchosal, M.A., Quadroni, M., et al. (2018). Ribonuclease inhibitor 1 regulates erythropoiesis by controlling GATA1 mRNA translation. J Clin Invest. Published online February 6, 2018. https://doi.org/10.1172/JCI94956.

Doulatov, S., Notta, F., Laurenti, E., and Dick, J.E. (2012). Hematopoiesis: a human perspective. Cell Stem Cell *10*, 120–136.

Engreitz, J.M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E.S., et al. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science *341*, 1237973.

Farrar, J.E., Quarello, P., Fisher, R., O'Brien, K.A., Aspesi, A., Parrella, S., Henson, A.L., Seidel, N.E., Atsidaftos, E., Prakash, S., et al. (2014). Exploiting pre-rRNA processing in Diamond Blackfan anemia gene discovery and diagnosis. Am. J. Hematol. *89*, 985–991.

Gautier, E.F., Ducamp, S., Leduc, M., Salnot, V., Guillonneau, F., Dussiot, M., Hale, J., Giarratana, M.C., Raimbault, A., Douay, L., et al. (2016). Comprehensive proteomic analysis of human erythropoiesis. Cell Rep. *16*, 1470–1484.

Gazda, H.T., Sheen, M.R., Vlachos, A., Choesmel, V., O'Donohue, M.F., Schneider, H., Darras, N., Hasman, C., Sieff, C.A., Newburger, P.E., et al. (2008). Ribosomal protein L5 and L11 mutations are associated with cleft palate and abnormal thumbs in Diamond-Blackfan anemia patients. Am. J. Hum. Genet. *83*, 769–780.

Giani, F.C., Fiorini, C., Wakabayashi, A., Ludwig, L.S., Salem, R.M., Jobaliya, C.D., Regan, S.N., Ulirsch, J.C., Liang, G., Steinberg-Shemer, O., et al. (2016). Targeted application of human genetic variation can improve red blood cell production from stem cells. Cell Stem Cell *18*, 73–78.

Gripp, K.W., Curry, C., Olney, A.H., Sandoval, C., Fisher, J., Chong, J.X., Pilchman, L., Sahraoui, R., Stabley, D.L., and Sol-Church, K.; UW Center for Mendelian Genomics (2014). Diamond-Blackfan anemia with mandibulofacial dystostosis is heterogeneous, including the novel DBA genes TSR2 and RPS28. Am. J. Med. Genet. A. *164A*, 2240–2249.

Henras, A.K., Plisson-Chastang, C., O'Donohue, M.F., Chakraborty, A., and Gleizes, P.E. (2015). An overview of pre-ribosomal RNA processing in eukaryotes. Wiley Interdiscip. Rev. RNA *6*, 225–242.

Hinnebusch, A.G., Ivanov, I.P., and Sonenberg, N. (2016). Translational control by 5′-untranslated regions of eukaryotic mRNAs. Science *352*, 1413–1416.

Hsieh, A.C., Liu, Y., Edlind, M.P., Ingolia, N.T., Janes, M.R., Sher, A., Shi, E.Y., Stumpf, C.R., Christensen, C., Bonham, M.J., et al. (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature *485*, 55–61.

Ingolia, N.T. (2016). Ribosome footprint profiling of translation throughout the genome. Cell *165*, 22–33.

Iskander, D., Psaila, B., Gerrard, G., Chaidos, A., En Foong, H., Harrington, Y., Karnik, L.C., Roberts, I., de la Fuente, J., and Karadimitris, A. (2015). Elucidation of the EP defect in Diamond-Blackfan anemia by characterization and prospective isolation of human EPs. Blood *125*, 2553–2557.

Jovanovic, M., Rooney, M.S., Mertins, P., Przybylski, D., Chevrier, N., Satija, R., Rodriguez, E.H., Fields, A.P., Schwartz, S., Raychowdhury, R., et al. (2015). Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. Science *347*, 1259038.

Khatter, H., Myasnikov, A.G., Natchiar, S.K., and Klaholz, B.P. (2015). Structure of the human 80S ribosome. Nature *520*, 640–645.

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.

Lee, W.Y., Weinberg, O.K., and Pinkus, G.S. (2017). GATA1 Is a Sensitive and Specific Nuclear Marker for Erythroid and Megakaryocytic Lineages. Am. J. Clin. Pathol. *147*, 420–426.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.

Li, J., Hale, J., Bhagia, P., Xue, F., Chen, L., Jaffray, J., Yan, H., Lane, J., Gallagher, P.G., Mohandas, N., et al. (2014). Isolation and transcriptome analyses of human erythroid progenitors: BFU-E and CFU-E. Blood *124*, 3636–3645.

Ludwig, L.S., Gazda, H.T., Eng, J.C., Eichhorn, S.W., Thiru, P., Ghazvinian, R., George, T.I., Gotlib, J.R., Beggs, A.H., Sieff, C.A., et al. (2014). Altered translation of GATA1 in Diamond-Blackfan anemia. Nat. Med. *20*, 748–753.

McCann, K.L., and Baserga, S.J. (2013). Genetics. Mysterious ribosomopathies. Science *341*, 849–850.

Mills, E.W., Wangen, J., Green, R., and Ingolia, N.T. (2016). Dynamic regulation of a ribosome rescue pathway in erythroid cells and platelets. Cell Rep. *17*, 1–10.

Mirabello, L., Khincha, P.P., Ellis, S.R., Giri, N., Brodie, S., Chandrasekharappa, S.C., Donovan, F.X., Zhou, W., Hicks, B.D., Boland, J.F., et al. (2017). Novel and known ribosomal causes of Diamond-Blackfan anaemia identified through comprehensive genomic characterisation. J. Med. Genet. *54*, 417–425.

Nathan, D.G., Clarke, B.J., Hillman, D.G., Alter, B.P., and Housman, D.E. (1978). Erythroid precursors in congenital hypoplastic (Diamond-Blackfan) anemia. J. Clin. Invest. *61*, 489–498.

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.

Ohene-Abuakwa, Y., Orfali, K.A., Marius, C., and Ball, S.E. (2005). Two-phase culture in Diamond Blackfan anemia: localization of erythroid defect. Blood *105*, 838–846.

Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell *132*, 631–644.

Paul, F., Arkin, Y., Giladi, A., Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., et al. (2015). Transcriptional heterogeneity and lineage commitment in myeloid progenitors. Cell *163*, 1663–1677.

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.

Preti, M., O'Donohue, M.F., Montel-Lehry, N., Bortolin-Cavaillé, M.L., Choesmel, V., and Gleizes, P.E. (2013). Gradual processing of the ITS1 from the nucleolus to the cytoplasm during synthesis of the human 18S rRNA. Nucleic Acids Res. *41*, 4709–4723.

Reschke, M., Clohessy, J.G., Seitzer, N., Goldstein, D.P., Breitkopf, S.B., Schmolze, D.B., Ala, U., Asara, J.M., Beck, A.H., and Pandolfi, P.P. (2013). Characterization and analysis of the composition and dynamics of the mammalian riboproteome. Cell Rep. *4*, 1276–1287.

Robledo, S., Idol, R.A., Crimmins, D.L., Ladenson, J.H., Mason, P.J., and Bessler, M. (2008). The role of human ribosomal proteins in the maturation of rRNA and ribosome production. RNA *14*, 1918–1929.

Roepcke, S., Zhi, D., Vingron, M., and Arndt, P.F. (2006). Identification of highly specific localized sequence motifs in human ribosomal protein gene promoters. Gene *365*, 48–56.

Sankaran, V.G., and Weiss, M.J. (2015). Anemia: progress in molecular mechanisms and therapies. Nat. Med. *21*, 221–230.

Sankaran, V.G., Ghazvinian, R., Do, R., Thiru, P., Vergilio, J.A., Beggs, A.H., Sieff, C.A., Orkin, S.H., Nathan, D.G., Lander, E.S., and Gazda, H.T. (2012). Exome sequencing identifies GATA1 mutations resulting in Diamond-Blackfan anemia. J. Clin. Invest. *122*, 2439–2443.

Schütz, S., Fischer, U., Altvater, M., Nerurkar, P., Peña, C., Gerber, M., Chang, Y., Caesar, S., Schubert, O.T., Schlenstedt, G., and Panse, V.G. (2014). A RanGTP-independent mechanism allows ribosomal protein nuclear import for ribosome assembly. eLife *3*, e03473.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Shah, P., Ding, Y., Niemczyk, M., Kudla, G., and Plotkin, J.B. (2013). Rate-limiting steps in yeast protein translation. Cell *153*, 1589–1601.

Shi, Z., Fujii, K., Kovary, K.M., Genuth, N.R., Rost, H.L., Teruel, M.N., and Barna, M. (2017). Heterogeneous ribosomes preferentially translate distinct subpools of mRNAs genome-wide. Mol. Cell *67*, 71–83.

Signer, R.A., Magee, J.A., Salic, A., and Morrison, S.J. (2014). Haematopoietic stem cells require a highly regulated protein synthesis rate. Nature *509*, 49–54.

Srivastava, D., and DeWitt, N. (2016). In vivo cellular reprogramming: the next generation. Cell *166*, 1386–1396.

Sung, M.K., Porras-Yakushi, T.R., Reitsma, J.M., Huber, F.M., Sweredoski, M.J., Hoelz, A., Hess, S., and Deshaies, R.J. (2016). A conserved quality-control pathway that mediates degradation of unassembled ribosomal proteins. eLife *5*, e19105.

Tanay, A., and Regev, A. (2017). Scaling single-cell genomics from phenomenology to mechanism. Nature *541*, 331–338.

Thompson, M.K., Rojas-Duran, M.F., Gangaramani, P., and Gilbert, W.V. (2016). The ribosomal protein Asc1/RACK1 is required for efficient translation of short mRNAs. eLife *5*, e11154.

Thoreen, C.C., Chantranupong, L., Keys, H.R., Wang, T., Gray, N.S., and Sabatini, D.M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. Nature *485*, 109–113.

Ulirsch, J.C., Nandakumar, S.K., Wang, L., Giani, F.C., Zhang, X., Rogov, P., Melnikov, A., McDonel, P., Do, R., Mikkelsen, T.S., and Sankaran, V.G. (2016). Systematic functional dissection of common genetic variation affecting red blood cell traits. Cell *165*, 1530–1545.

van Galen, P., Kreso, A., Mbong, N., Kent, D.G., Fitzmaurice, T., Chambers, J.E., Xie, S., Laurenti, E., Hermans, K., Eppert, K., et al. (2014). The unfolded protein response governs integrity of the haematopoietic stem-cell pool during stress. Nature *510*, 268–272.

Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D., et al. (2017). Human haematopoietic stem cell lineage commitment is a continuous process. Nat. Cell Biol. *19*, 271–281.

Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. Science *350*, 1096–1101.

Weiss, M.J., Yu, C., and Orkin, S.H. (1997). Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. Mol. Cell. Biol. *17*, 1642–1651.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat polyclonal anti-GATA1 (M-20)	Santa Cruz Biotechnology	Cat#: sc-1234; RRID: AB_2263157
Rabbit polyclonal anti-TSR2	Abcam	Cat#: ab155810; RRID: AB_2715561
Mouse monoclonal anti-RPS19 (WW-4)	Santa Cruz Biotechnology	Cat#: sc-100836; RRID: AB_1129199
Goat polyclonal anti-RPL5 (D-20)	Santa Cruz Biotechnology	Cat#: sc-103865; RRID: AB_2182039
Goat polyclonal anti-RPL11 (N-17)	Santa Cruz Biotechnology	Cat#: sc-25931; RRID: AB_2181298
Goat polyclonal anti-RPS20 (G-15)	Santa Cruz Biotechnology	Cat#: sc-55035; RRID: AB_2180344
Rabbit polyclonal anti-RPS24	Abcam	Cat#: ab102986; RRID: AB_10711571
Rabbit polyclonal anti-RPS26	Abcam	Cat#: ab104050; RRID: AB_10710999
Rabbit polyclonal anti-RPSA	Abcam	Cat#: ab137388; RRID: AB_2715562
Rabbit polyclonal anti-RPL28 (FL-137)	Santa Cruz Biotechnology	Cat#: sc-50362; RRID: AB_2181746
Rabbit polyclonal anti-EPOR (M-20)	Santa Cruz Biotechnology	Cat#: sc-697; RRID: AB_631468
Rabbit polyclonal anti-STAT5A (C-17)	Santa Cruz Biotechnology	Cat#: sc-835; RRID: AB_632446
Rabbit polyclonal anti-JAK2 (HR-758)	Santa Cruz Biotechnology	Cat#: sc-278; RRID: AB_631853
Mouse monoclonal anti-BYSTIN (A-10)	Santa Cruz Biotechnology	Cat#: sc-271722; RRID: AB_10707663
Goat polyclonal anti-LAMIN B (C-20)	Santa Cruz Biotechnology	Cat#: sc-6216; RRID: AB_648156
Mouse monoclonal anti-GAPDH (6C5)	Santa Cruz Biotechnology	Cat#: sc-32233; RRID: AB_627679
Mouse monoclonal anti-ACTB (AC-15)	Sigma Aldrich	Cat#: A1978; RRID: AB_476692
Donkey anti-mouse	Jackson ImmunoResearch	Cat#: 715-035-150; RRID: AB_2340770
Donkey anti-goat	Jackson ImmunoResearch	Cat#: 705-035-147; RRID: AB_2313587
Donkey anti-rabbit	Jackson ImmunoResearch	Cat#: 711-035-152; RRID: AB_10015282
APC anti-CD235a, clone HIR2	eBioscience	Cat#: 17-9987-42; RRID: AB_2043823
Pacific Blue anti-CD41a, clone HIP8	BioLegend	Cat#: 303714; RRID: AB_10696421
Pacific Blue anti-CD11b, clone ICRF44	BioLegend	Cat#: 301315; RRID: AB_493015
FITC anti-CD41a, clone HIP8	eBioscience	Cat#: 11-0419-42; RRID: AB_10718234
FITC anti-CD11b, clone ICRF44	BioLegend	Cat#: 301330; RRID: AB_2561703
Propidium Iodide	eBioscience	Cat#: 00-6990-50
APC Annexin V	BD PharMingen	Cat#: 550474
Alexa Fluor 488 anti-CD34, clone 581	BioLegend	Cat#: 343518; RRID: AB_1937203
Brilliant-Violet 421 anti-CD38, clone HB-7	BioLegend	Cat#: 356618; RRID: AB_2566231
PE anti-CD71, clone OKT9	eBioscience	Cat#: 12-0719-42; RRID: AB_10717077
FITC anti-CD235a, clone HIR2	BioLegend	Cat#: 306610; RRID: AB_756046
APC anti-TER-119, clone TER-119	eBioscience	Cat#: 17-5921-81; RRID: AB_469472
Rabbit monoclonal anti-GATA1, clone EP2819Y	Abcam	Cat#: ab76121; RRID: AB_1310256
Rabbit monoclonal IgG isotype, clone EPR25A	Abcam	Cat#: ab172730; RRID: AB_2687931
Alexa Fluor 647 Goat polyclonal anti-rabbit IgG (H+L)	Jackson ImmunoResearch	Cat#: 111-605-003; RRID: AB_2338072
Chemicals, Peptides, and Recombinant Proteins		
Dulbecco's Modified Eagle Medium-High Glucose (DMEM)	GIBCO	Cat#: 11965-118
Iscove's Modified Dulbecco's Medium (IMDM)	GIBCO	Cat#: 12440-061

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Roswell Park Memorial Institute Medium (RPMI) 1640	GIBCO	Cat#: 11875-119
Fetal Bovine Serum (FBS)	Atlanta Biologicals	Cat#: S11150
Human Holo-Transferrin	Sigma Aldrich	Cat#: T0665-1G
Penicillin-Streptomycin	GIBCO	Cat#: 15140-122
Human Serum, Type AB	Atlanta Biologicals	Cat#: S40110
Human Plasma, Type AB	Blood Bank at Boston Children's Hospital	N/A
Humulin R (Insulin)	Lilly	NDC 0002-8215-01
Heparin	Hospira	NDC 00409-2720-01
Epogen (recombinant erythropoietin)	Amgen	NDC 55513-267-10
Recombinant human stem cell factor (SCF)	PeproTech	Cat#: 300-07
Recombinant human interleukin-3 (IL3)	PeproTech	Cat#: 200-03
Opti-MEM	GIBCO	Cat#: 31985-062
StemSpan SFEM II medium	STEMCELL Technologies	Cat#: 09655
StemSpan CC100	STEMCELL Technologies	Cat#: 02690
1-Thioglycerol	Sigma Aldrich	Cat#: M6145
4X Laemmli Sample Buffer	Bio-Rad	Cat#: 161-0747
FuGENE 6 Transfection Reagent	Promega	Cat#: E2691
Dimethyl-sulfoxide (DMSO)	Sigma Aldrich	Cat#: D2438
Cycloheximide	Sigma Aldrich	Cat#: C7698
Polybrene Infection/Transfection reagent	Millipore	Cat#: TR-1003-G
RNasin Plus RNase Inhibitor	Promega	Cat#: N2615
SUPERase IN RNase Inhibitor	Ambion	Cat#: AM2696
cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail	Sigma Aldrich	Cat#: 11836170001
Protease Inhibitor Set	G-Biosciences	Cat#: 786-207
RNase I	Ambion	Cat#: AM2294
Trichloracetic acid	Sigma Aldrich	Cat#: T9159
Sodium deoxycholate	Sigma Aldrich	Cat#: 30970
RNA Polymerase I Inhibitor II, CX-5461	Millipore	Cat#: 509265
Critical Commercial Assays		
QuikChange site-directed mutagenesis Kit	Agilent Technologies	Cat#: 200518
RNeasy Plus Mini Kit	QIAGEN	Cat#: 74134
iScript cDNA Synthesis Kit	Bio-Rad	Cat#: 1708891
iQ SYBR Green Supermix	Bio-Rad	Cat#: 1708882
4-20% Mini-PROTEAN TGX Precast Protein Gels, 12 well	Bio-Rad	Cat#: 4561095
Clarity Western ECL Substrate	Bio-Rad	Cat#: 1705060
RIPA Lysis Buffer System	Santa Cruz Biotechnology	Cat#: sc-24948A
PARIS Kit	Ambion	Cat#: AM1921
Transcription Factor Buffer Set	BD PharMingen	Cat#: 562574
Direct-zol RNA MiniPrep Plus w/ TRI Reagent Kit	Zymo Research	Cat#: R2071
NEBNext Poly(A) mRNA Magnetic Isolation Module	New England Biolabs	Cat#: E7490
Truseq Ribo Profile (Mammalian) Kit	Illumina	Cat#: RPHMR12126

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Ribo-Zero Gold rRNA Removal Kit	Illumina	Cat#: MRZG126
illustra MicroSpin S-400 HR Columns	GE Healthcare Life Sciences	Cat#: 27-5140-01
Zero Blunt TOPO PCR Cloning Kit	invitrogen	Cat#: 450031
Deposited Data		
Whole-exome sequencing data	https://www.ncbi.nlm.nih.gov/gap	dbGAP accession: phs000474.v2.p1
Raw mass spectrometry data	This study	MassIVE: MSV000080283
RNA-seq and ribosome profiling data	This study	GEO: GSE89183
Experimental Models: Cell Lines		
Human CD34$^+$ hematopoietic stem and progenitor cells, adult	Fred Hutchinson Cancer Research Center	N/A
Human CD34$^+$ hematopoietic stem and progenitor cells, adult	Division of Hematology/ Oncology Flow Cytometry Research Facility at Boston Children's Hospital	N/A
K562 cells	ATCC	Cat#: CCL-243
G1E cells	Weiss et al., 1997	N/A
Experimental Models: Strains		
Yeast: P_{GAL1}-*TSR2*	Schütz et al., 2014	N/A
Sequence Based Reagents		
shTSR2-1_TRCN0000172642: CCGGG AGGTCACAGCTACGAATGATCTCGAG ATCATTCGTAGCTGTGACCTCTTTTTG	Sigma Aldrich	N/A
shTSR2-2_TRCN0000344162: CCGGAG GATTACTTCATGCGCAATGCTCGAGCA TTGCGCATGAAGTAATCCTTTTTG	Sigma Aldrich	N/A
shRPS19-1_TRCN0000074913: CCGGC TACGATGAGAACTGGTTCTACTCGAGT AGAACCAGTTCTCATCGTAGTTTTTG	Sigma Aldrich	N/A
shRPS19-2_TRCN0000074916: CCGGG CTTGCTCCCTACGATGAGAACTCGAGT TCTCATCGTAGGGAGCAAGCTTTTTG	Sigma Aldrich	N/A
shRPL5-1_TRCN0000074994: CCGGG TTCGTGTGACAAACAGAGATCTCGAG ATCTCTGTTTGTCACACGAACTTTTTG	Sigma Aldrich	N/A
shRPL5-2_TRCN0000074997: CCGGC CCTCACAGTACCAAACGATTCTCGAG AATCGTTTGGTACTGTGAGGGTTTTTG	Sigma Aldrich	N/A
shRPS24-1_TRCN0000117550: CCGGC GCAAGAACAGAATGAAGAAACTCGAG TTTCTTCATTCTGTTCTTGCGTTTTTG	Sigma Aldrich	N/A
shRPS24-2_TRCN0000117551: CCGGG ATTTATGATTCCCTGGATTACTCGAGTA ATCCAGGGAATCATAAATCTTTTTG	Sigma Aldrich	N/A
shRPL11-1_TRCN0000117712: CCGGG CGGGAGTATGAGTTAAGAAACTCGAGT TTCTTAACTCATACTCCCGCTTTTTG	Sigma Aldrich	N/A
shRPL11-2_TRCN0000117713: CCGGC CGCAAACTCTGTCTCAACATCTCGAGA TGTTGAGACAGAGTTTGCGGTTTTTG	Sigma Aldrich	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
See Table S6 for all primers used for RT-qPCRs.	This study	N/A
See Table S5 for cloned 5′UTR and GATA1 coding sequences.	This study	N/A
Recombinant DNA		
pRS425-*hTSR2*	Schütz et al., 2014	N/A
pRS425-*hTsr2E64G*	This study	N/A
HMD-GATA1	Ludwig et al., 2014	N/A
Software and Algorithms		
Image Lab Version 5.2.1	Bio-Rad	http://www.bio-rad.com/en-cn/product/image-lab-software?ID=KRE6P5E8Z
FlowJo 10.0.7	FlowJo	https://www.flowjo.com/solutions/flowjo
GraphPad Prism 7	Graphpad Software Inc	https://www.graphpad.com/scientific-software/prism/
R version 3.2	The R Foundation	https://www.r-project.org
ExAC	Lek et al., 2016	http://exac.broadinstitute.org
Picard tools	Broad Institute	https://broadinstitute.github.io/picard/
Spectrum Mill MS Proteomics Workbench v6.0 pre-release software package	Agilent Technologies	http://proteomics.broadinstitute.org
FastQC	Babraham Bioinformatics	http://www.bioinformatics.babraham.ac.uk/projects/fastqc
CellProfiler	Carpenter et al., 2006	http://cellprofiler.org

CONTACT FOR REAGENT AND RESOURCE SHARING

Requests for further information or reagents may be directed to the Lead Contact, Vijay G. Sankaran (sankaran@broadinstitute.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Primary Cell Culture

$CD34^+$ cells were obtained from magnetically sorted mononuclear samples of G-CSF–mobilized peripheral blood from donors and were frozen after isolation. Cells were obtained from the Fred Hutchinson Cancer Research Center, Seattle, USA or the Division of Hematology/Oncology Flow Cytometry Research Facility at Boston Children's Hospital. Cells were thawed and washed into PBS with 1% human AB serum (Atlanta Biologicals), pelleted and then seeded in differentiation medium containing IMDM with 2% human AB plasma, 3% human AB serum, 1% P/S, 200 μg/mL holo-transferrin, 10 ng/mL SCF (PeproTech, Inc.), 1 ng/mL IL-3 (PeproTech, Inc.) and 3 U/mL erythropoietin (EPO) (Amgen). Where an expansion phase is indicated, $CD34^+$ cells were cultured in StemSpan SFEM II medium (STEMCELL Technologies) supplemented by 1X CC100 (containing FLT3 ligand, stem cell factor (SCF), IL-3, and IL-6, STEMCELL Technologies) for 5 days prior to differentiation. Cells were maintained at a density between 0.1×10^6 and 0.5×10^6 cells per milliliter, with medium changes every other day as necessary. Cells were incubated at 37°C with 5% CO_2.

293T and K562 Cell Culture

293T cells (ATCC) were maintained in DMEM with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (P/S). K562 human erythroid cells (ATCC) were maintained at a density between 0.1×10^6 and 1×10^6 cells per milliliter in RPMI 1640 medium supplemented with 10% FBS and 1% P/S. Cells were incubated at 37°C with 5% CO_2.

G1E Cell Culture

G1E cells (Weiss et al., 1997) were cultured in IMDM with 15% FCS, 1% P/S., 4.5×10^{-5} M Monothioglycerol (MTG), 50ng/ml SCF and 2 U/ml EPO at a density between 0.1×10^6 and 1×10^6 cells per ml, with medium changes every day as necessary. Cells were incubated at 37°C with 5% CO_2.

METHOD DETAILS

Lentiviral Vectors and Infection

The shRNA constructs targeting human *TSR2* (shTSR2-1 and shTSR2-2, RefSeqID NM_058163), human *RPS19* (shRPS19-1 and shRPS19-2, RefSeqID NM_001022), human *RPS24* (shRPS24-1 and RPS24-2, RefSeq ID NM_001026), human *RPL5* (shRPL5-1 and RPL5-2 RefSeq ID NM_000969) and human *RPL11* (shRPL11-1 and RPL11-2 RefSeq ID NM_000975) were obtained from the Mission shRNA collection (Sigma-Aldrich). The sequences of the shRNAs used in this study are listed in the Key Resources Table.

The lentiviral vectors pLKO-GFP and pLKO.1 targeting Luciferase (shLuc) (Genetic Pertubation Platform of the Broad Institute of MIT and Harvard) were used as controls. Rescue experiments were performed as described previously (Ludwig et al., 2014) by co-transduction of human erythroid cells with shRNAs targeting TSR2 and either the HMD control or HMD-GATA1, which contain the respective cDNAs. Double-transduced cells were identified by puromycin selection and GFP expression driven by an IRES-GFP in the HMD vector.

For lentivirus production, 293T cells were transfected with pVSV-G and pDelta8.9 using FuGene 6 reagent (Promega) according to the manufacturer's protocol. The medium was changed the day after transfection to the appropriate culture medium. After 30 h, viral supernatant was collected and filtered using a 0.45 μm filter immediately before infection of primary hematopoietic or K562 cells in a 6-well plate at a density of 200,000–500,000 cells per well in the presence of 8 μg/ml polybrene (Millipore). The cells were spun at 2,000 rpm for 90 min at 22°C and left in viral supernatant overnight. The medium was replaced the morning after infection. Puromycin selection of infected cells was started 36 h after infection with 1 μg/ml for primary hematopoietic cells or 2 μg/ml for K562 cells. Infection efficiency was between 50%–80% for primary hematopoietic cells and > 95% for K562 cells as assessed by flow cytometry of pLKO-GFP infected cells.

Yeast Strains and Plasmids

Preparation of media, yeast transformations and genetic manipulations were performed according to established procedures (Schütz et al., 2014). Plasmids used in this study are listed in Table S6. Details of plasmid construction will be provided upon request. All recombinant DNA techniques were performed according to established procedures using *E. coli* XL1 blue cells for cloning and plasmid propagation. Point mutations in human TSR2 were generated using the QuikChange site-directed mutagenesis kit (Agilent Technologies). All cloned DNA fragments and mutagenized plasmids were verified by sequencing. The PGAL1-TSR2 strain transformed with indicated plasmids in Figure 1 was spotted in 10-fold dilutions on selective glucose containing plates and grown at indicated temperatures for 3-7 days.

Quantitative RT-PCR

Isolation of RNA was performed using the RNeasy Plus Mini Kit (QIAGEN). An on-column DNase (QIAGEN) digestion was performed according to the manufacturer's instructions. RNA was quantified by a NanoDrop spectrophotometer (Thermo Scientific). Reverse transcription was carried out using the iScript cDNA synthesis kit (Bio-Rad). Real-time PCR was performed using the CFX96 Real-time PCR detection system (Bio-Rad) and iQ SYBR® Green Supermix (Bio-Rad). Quantification was performed using the comparative C_T method. Normalization was performed using β-actin mRNA as a standard. The primers used for quantitative RT-PCR are listed in Table S6.

Western Blotting

Cells were harvested 5 days post-infection or at 72 h of treatment with the polymerase I inhibitor CX-5461 (Millipore), washed twice in PBS, resuspended in RIPA lysis buffer (50 mM Tris-HCl at pH 7.4, 150 mM NaCl, 0.1% SDS, 1% NP-40, 0.25% sodium deoxycholate, 1 mM DTT) supplemented with 1 × Complete Protease Inhibitor Cocktail (Roche) and incubated for 30 min on ice. After centrifugation at 15,000 rpm for 10 min at 4°C to remove cellular debris, the supernatant was transferred to a new tube, supplemented with Laemmli sample buffer (Bio-Rad) and incubated for 10 min at 90°C. Equal amounts of proteins were separated by SDS gel electrophoresis using the Mini-PROTEAN® TGX gel system (Bio-Rad) and Tris/glycine/SDS running buffer. Subsequently, proteins were transferred onto a PVDF membrane (Millipore) using Tris/glycine transfer buffer. Membranes were blocked with 3% BSA-PBST for 1 h and probed with GATA1 goat polyclonal antibody (M-20, sc-1234, Santa Cruz Biotechnology) at a 1:500 dilution, TSR2 rabbit polyclonal antibody (ab155810, Abcam) at a 1:1,000 dilution, RPS19 mouse monoclonal antibody (WW-4, sc-100836, Santa Cruz Biotechnology) at a 1:500 dilution, RPL5 goat polyclonal (D-20, sc-103865, Santa Cruz Biotechnology) at a 1:500 dilution, RPL11 goat polyclonal (N-17, sc-25931, Santa Cruz Biotechnology) at a 1:500 dilution, RPS20 goat polyclonal (G-15, sc-55035, Santa Cruz Biotechnology) at a 1:500 dilution, RPS24 rabbit polyclonal (ab102986, Abcam) at a 1:1,000 dilution, RPS26 rabbit polyclonal (ab104050, Abcam) at a 1:1,000 dilution, RPSA rabbit polyclonal (ab137388, Abcam) at a 1:1,000 dilution, RPL28 rabbit polyclonal (FL-137, sc-50362, Santa Cruz Biotechnology) at a 1:1,000 dilution, EPOR rabbit polyclonal (M-20, sc-697, Santa Cruz Biotechnology) at a 1:500 dilution, STAT5A rabbit polyclonal (C-17, sc-835, Santa Cruz Biotechnology) at a 1:500 dilution, JAK2 rabbit polyclonal (HR-758, sc-278, Santa Cruz Biotechnology) at a 1:500 dilution, Bystin mouse monoclonal (A-10, sc-271722, Santa Cruz Biotechnology) at a 1:1,000 dilution, Lamin B goat polyclonal (C-20, sc-6216, Santa Cruz Biotechnology) at a 1:500 dilution, ACTB mouse monoclonal (AC-15, A1978, Sigma Aldrich) at a 1:10,000 dilution or GAPDH mouse monoclonal antibody (6C5, sc-32233, Santa Cruz Biotechnology) at a 1:1,000 dilution in 3% BSA-PBST over-night at 4°C. Membranes were washed

four times with PBST, incubated with donkey anti-mouse, anti-goat or anti-rabbit peroxidase-coupled secondary antibodies (715-035-150, 705-035-147 or 711-035-152, respectively; Jackson ImmunoResearch) at a 1:10,000 to 1:20,000 dilution in 3% BSA-PBST for 1 h at room temperature, washed three times with PBST and incubated for 5 minutes with Clarity western ECL substrate (Bio-Rad). Proteins were visualized by using the ChemiDoc Touch Imaging System (Bio-Rad) or by exposure to scientific imaging film (GE Healthcare Life Sciences). Band intensities were determined with Image Lab (Bio-Rad). Where indicated, separation of nuclear and cytoplasmic fractions was performed with the PARIS^{TM} Kit (Ambion).

Flow Cytometry Analysis and Apoptosis Detection

For flow cytometry analysis, *in vitro* cultured hematopoietic cells were washed in PBS and stained with propidium iodide (PI), 1:20 APC-conjugated CD235a (glycophorin A, clone HIR2, eBioscience), 1:20 Pacific Blue-conjugated CD41a (HIP8, BioLegend), 1:20 Pacific Blue-conjugated CD11b (ICRF44, BioLegend) or 1:25 APC-conjugated TER-119 (TER-119, eBioscience). For apoptosis analysis, the Annexin V-APC staining kit was used according to the manufacturer's instructions (550474, BD PharMingen). FACS analysis was conducted on a BD Bioscience Canto II flow cytometer. Data were analyzed using FlowJo 10.0.7 (TreeStar).

Intracellular GATA1 Staining

Uncultured, frozen cells from healthy individuals' and DBA patients' bone marrow specimens were recovered and stained for comparison of GATA1 protein expression in HSPC populations. Primary human adult HSPCs (from mobilized peripheral blood derived from G-CSF treated donors) were collected and stained at different time points of an *in vitro* culture system with expansion and differentiation phases to assess GATA1 expression at different stages of erythroid differentiation from unperturbed HSPCs. For each experiment, K562 and 293T cells were used as internal positive and negative controls, respectively. Cells were rinsed with 0.5% BSA in 1X PBS and stained for surface markers with CD34 Alexa488 (clone 581, BioLegend) and CD38 BV421 (clone HB7, BioLegend), or for CD71 PE (clone OKT9, eBioscience) and CD235a FITC (clone HIR2, BioLegend). Cells were then fixed, permeabilized and stained for GATA1 according to the BD PharmingenTM Transcription Factor Buffer Set protocol (BD PharMingen). 1:100 GATA1 rabbit monoclonal antibody EP2819Y (Abcam) or 1:200 rabbit monoclonal IgG isotype control were used as primary antibodies and polyclonal goat anti-rabbit IgG (H+L) Alexa647 conjugate (Jackson) was used as secondary antibody. Cells were run on BD Accuri C6 or BD Fortessa flow cytometers. Data were analyzed using FlowJo 10.2 (TreeStar).

rRNA Processing Examination

Northern blot analysis was done as described previously (Farrar et al., 2014). Bioanalyzer traces were obtained on an Agilent 2100 system with RNA Pico 6000 chips, sample processing was done according to the manufacturer's instructions.

Polysome Profiling

Cells were incubated with 100 μg/ml of cycloheximide (Sigma Aldrich) for 5 min at 37°C, washed twice with ice-cold PBS containing 100 μg/ml of cycloheximide and lysed in 10 mM Tris-HCl (pH 7.4), 5 mM $MgCl_2$, 100 mM KCl, 1% Triton X-100, 3 mM DTT, 100 μg/ml cycloheximide, 500 U/ml RNasin (Promega) and 1 × Complete Protease Inhibitor, EDTA-free (Roche) as well as 1x Protease Inhibitor Set (without EDTA) (G-Biosciences). Polysomes were separated on a 10%–50% (or 10%–45%) linear sucrose gradient containing 20 mM Tris-HCl (pH 7.4), 5 mM $MgCl_2$, 100 mM KCl, 3 mM DTT, 100 μg/ml cycloheximide and 20 U/ml SUPERase• In RNase Inhibitor (Ambion) and centrifuged at 36,000 rpm for 2 h in a SW41 rotor in an L8-80M ultracentrifuge (Beckman Coulter). For mass spectrometry samples, gradients were fractionated using a Biocomp Gradient Station fractionator. Absorbance at 254 nm was used to visualize the gradients using an Econo UV monitor (Bio-Rad). Further processing for mass spectrometry analyses is described below.

Mass Spectrometry

Collected fractions for monosomes (a single ribosome), light polysomes (2-4 ribosomes) and heavy polysomes ($\geq$5 ribosomes) from K562 cells with indicated knockdown were pooled, respectively. Proteins from respective fractions were precipitated with deoxycholate-trichloracetic acid as described previously (Reschke et al., 2013), protein pellets were resuspended in 50 mM Tris HCl buffer containing 8 M Urea at pH8. Protein concentrations of the samples were estimated by BCA protein assay (ThermoFisher Scientific). Samples were reduced with 20 mM dithiothreitol at 37°C for 30 min, and alkylated with 50 mM iodoacetamide at room temperature in the dark for 30 min. Urea concentration was diluted to 2 M with 50 mM Tris HCl pH8 prior to Lys-C digestion (Wako) at 1:50 (w:w) enzyme to substrate ratio at 30°C for 2 h with mixing on the shaker at 850 rpm. Urea was further diluted to less than 1 M prior to overnight digestion with trypsin (Promega) with 1:50 (w:w) enzyme to substrate ratio at 37°C with shaking at 850 rpm. Digestion was terminated with formic acid to a final concentration of 1%. The digests were desalted on vacuum manifold using Oasis HLB 1cc (30 mg) reversed phase cartridges (Waters) with 0.1% formic acid/water and 0.1% formic acid/80% acetonitrile as buffers A and B, respectively. Briefly, cartridges were conditioned with 3 × 500 μL buffer B followed by equilibration with 4 × 500 μL buffer A. After loading the digests at a reduced flow rate, they were washed with 3 × 750 μL buffer A and eluted with 3 × 500 μL buffer B. Eluates were frozen and dried by vacuum centrifugation. Digests were reconstituted in 100 μL of 0.1% formic acid /3% acetonitrile and post-digestion concentrations were determined by NanoDrop 2000 (ThermoFisher Scientific). Based on the post-digestion concentration, 30 μg aliquots were prepared, dried to dryness by vacuum centrifugation and stored at −80°C. A pooled reference sample was created

by mixing equal amounts of the monosome, light and heavy polysome samples from both replicates of the shLuc control cell line and aliquoted at 30 μg, dried to dryness by vacuum centrifugation and stored at −80°C.

TMT ten-plex reagent (ThermoFisher Scientific) was used for isobaric labeling of samples. Sample labeling was designed so that duplicate samples from all four cell lines representing a given sucrose gradient fraction (monosomes, light polysomes, heavy polysomes) were contained within the same TMT ten-plex experiment with the reference sample included as the 9th channel in all three TMT ten-plex experiments. The 10th channel was omitted from the experiment. Table S7 summarizes TMT reagent channel line-up for all the samples.

Thirty microgram dried aliquot of each sample was labeled with TMT ten-plex reagent following manufacturer's instructions (ThermoFisher Scentific). Samples were reconstituted in 30 μL 50 mM HEPES buffer. 800 μg of each TMT reagent was reconstituted in 41 μL acetonitrile and 12.3 μL of the resulting solution was added to each sample, mixed and incubated at room temperature for 1 h with shaking at 850 rpm. Three microliters of each sample was used to check label incorporation by LC-MS/MS prior to quenching the reaction. Once satisfied with labeling efficiency (> 95% label incorporation) the reactions were quenched by adding 2.4 μL of 5% hydroxylamine to a 0.08 μg/μL concentration and incubated at room temperature for 15 min with shaking. Labeled samples representing each fraction type along with the pooled reference control were mixed together, dried down and desalted using Oasis HLB 1cc (30 mg) reversed phase cartridges as described above. Eluates were frozen, dried to dryness, and stored at −80°C.

Samples were reconstituted in 0.1% formic acid/3% Acetonitrile at 1 μg/μL concentration and 1 μL of it was analyzed on Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nLC 1000 UHPLC system (Proxeon, Thermo Fisher Scientific). Chromatography was performed on a 75 μm ID picofrit column (New Objective) packed in house with Reprosil-Pur C18 AQ 1.9 μm beads (Dr. Maisch, GmbH) to a length of 20 cm. Columns were heated to 50 °C using column heater sleeves (Phoenix-ST). Mobile phases consisted of 0.1% formic acid/3% acetonitrile as solvent A, and 0.1% formic acid/90% acetonitrile as solvent B. Peptides were eluted at 200 nL/min with a gradient of 6 to 35% B in 150 min, 35 to 60% B in 8 min, 60 to 90% B in 3 min, hold at 90% B for 10 min, 90% B to 50% B in 1 min, followed by isocratic hold at 50% B for 10 min. A single Orbitrap MS scan from 300 to 1800 m/z at a resolution of 70,000 with AGC set at 3e6 was followed by up to 12 ms/ms scans at a resolution of 35,000 with AGC set at 5e4. MS/MS spectra were collected with normalized collision energy of 29 and isolation width of 1.6 amu with isolation offset set to 0.3 amu. Dynamic exclusion was set to 20 s, and peptide match was set to preferred. Data analysis is described below.

RNA Polymerase I Inhibition

Human CD34$^+$ cells were cultured in erythroid differentiation medium as described above. Treatment with the RNA polymerase I inhibitor CX-5461 (Millipore) was started on day 3 of differentiation. Flow cytometry analysis was performed at 72 hours of CX-5461 treatment, with propidium iodide (eBioscience), 1:40 APC-conjugated CD235a (glycophorin A, clone HIR2, eBioscience), 1:40 FITC-conjugated CD41a (clone HIP8, eBioscience) and 1:40 FITC-conjugated CD11b (clone ICRF44, BioLegend). Samples were run on a BD LSRFortessa. Protein lysates for western blot analyses were collected at 72 hours of CX-5461 treatment. The western blot procedure is described above.

Ribosome Profiling

Lysates were prepared as described under polysome profiling and partitioned for either ribosome footprint profiling or mRNA sequencing. Total RNA was extracted with the Direct-zol RNA MiniPrep Plus w/ TRI Reagent® Kit (Zymo Research) according to the manufacturer's instructions. Total mRNA was poly-A selected using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (New England Biolabs) according to the manufacturer's instructions. mRNA seq libraries were generated as described previously (Engreitz et al., 2013). Ribosome footprinting and subsequent library preparation of ribosome protected RNA fragments (RPFs) was performed with the Truseq Ribo Profile (Mammalian) Kit (Illumina) according to the manufacturer's protocol. rRNA removal was performed by using the Ribo-Zero Gold rRNA Removal Kit (Illumina). RNase I (Ambion) digestion was done at a concentration of 2.5 U/μl lysate. RPFs were purified with MicroSpin S-400 columns (GE Healthcare Life Sciences). All libraries were sequenced on a HiSeq 2500 system (Illumina).

5′UTR-GATA1 Construct Cloning

5′ UTRs were defined from CD34$^+$ HSPC CAGE data. For cloning, the RUNX1 and GATA1 5′UTR-GATA1 constructs were synthesized by Integrated DNA Technologies (IDT). Alternatively, the GATA1 coding region was synthesized by IDT, and joined to PCR amplified ETV6 and LMO2-5′UTR fragments by overlap PCR and TOPO cloned (Zero Blunt TOPO PCR Cloning Kit, Invitrogen). Finally, all fragments were cloned into the U6_optisgRNA_modEF1s_p2A_GFP vector using BamHI and XhoI restriction sites. All constructs were verified by Sanger sequencing. Relevant construct sequences are shown in Table S5.

QUANTIFICATION AND STATISTICAL ANALYSIS

Whole Exome Sequencing

The cousins described in this manuscript underwent whole exome sequencing at the Broad Institute (dbGAP accession phs000474.v2.p1). In this study, whole exome sequencing and variant calling was performed as previously reported (Sankaran et al., 2012). Coverage across protein coding regions was calculated using Picard tools (Table S2). Variant Effect Predictor v83

(https://www.ensembl.org/info/docs/tools/vep/index.html) and the dbNSFP database v3.1 (https://sites.google.com/site/jpopgen/dbNSFP) were used to annotate the variant call file (VCF). We did not identify any rare (defined as 0.01% allele frequency in ExAC v0.3) (Lek et al., 2016) damaging (missense or loss of function) mutations in any of the known DBA genes (*RPS19 (revised nomenclature (*http://www.bangroup.ethz.ch/research/nomenclature-of-ribosomal-proteins.html*): eS19), RPL5 (uL18), RPL11 (uL5), RPL35A (eL33), RPL35 (uL29), RPS26 (eS26), RPS24 (eS24), RPS17 (eS17), RPS7 (eS7), RPS10 (eS10), RPL26 (uL24), RPS29 (uS14), RPS28 (eS28), RPS27 (eS27), RPL27 (eL27), RPL15 (eL15), RPL31 (eL31), RPL18 (eL18), GATA1*) or in any other ribosome protein coding genes that fit the predicted dominant or X-linked inheritance pattern. We thus investigated all genes for rare and predicted damaging mutations that fit either of these inheritance patterns (Table S1). Subsequently, we identified chrX:54469851:A > G in *TSR2* as the most likely candidate and verified this mutation by Sanger sequencing.

ExAC Gene Constraint Analyses

The ExAC v0.3 database, containing allele frequencies from whole exome sequencing for 60,706 unrelated individuals lacking Mendelian pediatric disease, has been used to estimate the probability that any single gene is intolerant to LoF mutations (known as pLI) (Lek et al., 2016). We compared the distribution of probabilities for a random sample of all genes (for ease of plotting) to RP genes and known DBA genes. Mann-Whitney-U tests were used to determine if there were significant differences in pLI between groups.

Analysis of Mass Spectrometry Data

Data extraction and searching was done using Spectrum Mill MS Proteomics Workbench v6.0 pre-release software package (Agilent Technologies). All extracted spectra were searched against a UniProt database containing human reference proteome sequences. Search was done using parent and fragment mass tolerance of 20ppm, and enzyme specificity set to trypsin allow P with 4 missed cleavages. Cysteine carbamidomethylation and TMT labeling at lysine and N-termini were set as fixed modifications. Allowed variable modifications were acetylation of protein N-termini, oxidized methionine, deamidation of asparagine, pyro-glutamic acid at peptide N-terminal glutamine, and pyro-carbamidomethylation at peptide N-terminal cysteine. Autovalidation was performed at peptide level with set FDR of less than 0.8 for charges 2 to 4, and less than 0.4 for charge 5 followed by protein level with set protein FDR of 0. Subgroup specific grouping of proteins was used for generating final protein table for each of the TMT experiments, which ensures that only peptides specific to a particular isoform are used for quantitation. Reporter ion intensities were corrected for isotopic impurities in the Spectrum Mill protein/peptide summary module using the static correction method and correction factors obtained from the reagent manufacturer's certificate of analysis (https://www.thermofisher.com/order/catalog/product/90406) for lot number QE214905A.

Only ribosomal proteins with 2 or more distinct peptides were used for further analysis of the data. For each TMT experiment representing one of the ribosomal fractions the normalized expression for protein *i* in TMT channel *j* is calculated using the following equation:

$$\frac{Ii}{\sum_{K \in RP} IK} \times \frac{Tj}{\sum_{t=1}^{9} Tt} \times \frac{\#O}{\#A}$$

Where *I* = protein precursor intensity, *i* = protein, *RP* = ribosomal proteins, *T* = TMT channel abundance for a given protein, *j* = TMT channel, #*O* = number of observed peptides for a protein, #*A* = number of theoretical peptides for a protein. The first term represents fractional precursor intensity over all observed ribosomal proteins; the second term is the fractional TMT reporter intensity and the final term adjusts for protein length.

All normalized values were then $\log_2$ transformed and median centered for each TMT channel. These values were used for all subsequent statistical analyses. Standard linear regression was performed between groups (shLuc, shRPL5, shRPS19, shTSR2) for different fractions (M, LP, HP) and for different subunits (80S, 60S, 40S). Linear fits and Pearson correlation coefficients are reported. Studentized, or jack-knifed, residuals were calculated in R using the studres() function in the MASS R package.

The ribosome-associated proteins were analyzed by identifying proteins that were similarly abundant as RPs in fractions of actively translating polysomes (LP and HP) in controls or cells with ribosomal perturbations. To do so, we have plotted intensity/density profiles for the HP/LP samples, in which we noted that the density (a smoothed histogram) was bimodal. We then used a mixture model, which essentially clustered the proteins into two groups - one RP-like and the other containing the remaining proteins. If a protein was in this RP-like cluster for any HP/LP sample, it was included in the analysis. In total, we identified 227 proteins (excluding the RPs) that fell into this cluster.

Analysis of RNA and Ribosome Profiling Libraries

Raw reads were trimmed using cutadapt with the options "-q 5 -m 20–discard-untrimmed -a AGATCGGAAGAGCACACGTCTG" (https://cutadapt.readthedocs.io/en/stable/). Bowtie2 was then used to align trimmed reads to rRNA, tRNA, and abundant noncoding RNAs (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml). FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) was used to determine that adapters and other sequences had been removed and to calculate the fragment length distribution of RPFs. The remaining reads were then aligned to the human hg19 genome build allowing for junctions based upon ENSEMBL transcripts using Tophat with the options "–no-novel-juncs–library-type fr-unstranded" (http://ccb.jhu.edu/software/tophat/index.shtml).

Non-uniquely mapping reads were excluded using Samtools (http://samtools.sourceforge.net). RSeQC was used to determine the percentage of reads mapping to 5′ UTRs, CDS, and 3′ UTRs (http://rseqc.sourceforge.net). Triplet periodicity was assessed using RibORF (https://personal.broadinstitute.org/zheji/software/RibORF.html). For RNA-seq, genes were quantified either using Cuffquant and Cuffnorm with the option "-max-bundle-frags 20000000" or using HTSeq-count in intersection-strict mode. Fragments per kilobase per million (FPKM) were subsequently transformed to transcripts per million (TPM). For RPFs, reads between 26 and 34 nucleotides in length were quantified in the CDSs of protein coding genes using HTSeq-count. Reads mapping to less than 45 nucleotides from the start codon or 15 nucleotides from the stop codon were not included in order to reduce read biases in the 5′ and 3′ ends of CDSs. To determine differentially expressed genes between control and RPH or TSR2 suppression conditions, we used a negative binomial model (mean and variance of distribution estimated in DESeq2) (http://bioconductor.org/packages/devel/bioc/vignettes/DESeq2/inst/doc/DESeq2.html). To determine differentially translated (e.g., changes in TE) genes, we used Xtail (https://github.com/xryanglab/xtail), which first uses the negative binomial distribution to estimate either (1) the $\log_2$ fold changes separately for mRNAs and RPFs between conditions (i.e., ΔmRNA and ΔRPF) or (2) the $\log_2$ fold changes for mRNA to RPF within conditions (i.e., $TE_{control}$ and TE_{RPH}), and then estimates a discrete joint probability distribution of either (1) ΔmRNA and ΔRPF or (2) $TE_{control}$ and TE_{RPH}. Testing of differential translation (i.e., ΔTE in both cases) was then performed, the least significant result of the two methods was kept. The Benjamini–Hochberg FDR was used to control for multiple testing. Only genes with > 150 mRNA counts and > 90 RPF counts were analyzed in order to obtain more stable estimates of ΔTE. Gene set enrichment analysis (GSEA) was used with the "Preranked" option and 10,000 permutations for ΔTE or ΔmRNA. An erythroid gene set was derived by taking all genes that were > 4 $\log_2$ fold upregulated between CD34+ and pro-erythroblast stages of normal human erythropoiesis. In addition, BIOCARTA, KEGG, and REACTOME canonical pathways were investigated.

Re-Annotation of 5′ UTRs

Because the TSS of a gene can vary between cell types and is often misannotated, we used cap analysis of gene expression (CAGE) data from K562 cells to define empirical TSS locations at 10-bp resolution using a heuristic algorithm. Four replicates of CAGE data (aligned BAM files CNhs12334.10824-111C5, CNhs12335.10825-111C6, CNhs12336.10826-111C7, and CNhs11250.10454-106G4) were downloaded from the FANTOM project (Arner et al., 2015) and merged using samtools. Each Ensembl gene (+/− 1 kb around the annotated ends of the gene) was scanned at 20-bp resolution to find the 100-bp window with the most number of CAGE reads, considering strand. Additional windows were chosen until either the windows either contained 80% of the total reads overlapping the gene or until these windows, upon merging of overlapping regions, contained 500bp of sequence. The top region was further scanned to find the 10-bp window with the most number of reads. This 10-bp window was defined as the empirical TSS. Next, empirical 5′ UTRs were determined by overlapping empirical TSSs with annotated ENSEMBL 5′ UTR positions for each transcript. When the empirical TSS fell within the annotated 5′ UTR, the 5′ UTR was shortened to start at the empirical TSS. When the empirical TSS was upstream of the annotated 5′ UTR, the 5′ UTR was extended to the empirical TSS. In all cases, the shortest 5′ UTR for a gene across all transcripts was taken, genes without empirical TSSs were excluded, and only genes with empirical 5′ UTRs < 500 nucleotides were included. Manual investigation of genes with 5′ UTRs > 500 nucleotides revealed that the majority of these were false positives that often had weak CAGE signal and/ or poor initial annotations. Additionally, $CD34^+$ HSPC CAGE data was downloaded from the ENCODE project (ENCFF000TTH.bam). For the 36 hematopoietic TFs investigated in $CD34^+$ CAGE, single nucleotide TSSs were identified based upon the strongest CAGE signal at any single nucleotide.

Analysis of Features for Association with ΔTE

A number of features were investigated for differences between RPH-sensitive and unchanged genes. The complexity of the empirical 5′ UTR secondary structure was determined using RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). Gene expression during erythropoiesis was performed by An et al. (2014) and Li et al. (2014) and processed as previously described (Ulirsch et al., 2016). Protein abundances for erythroid progenitors ("prog2") were obtained from Gautier et al. (2016). ORF lengths were calculated for the most abundant transcript (determined by highest TPM from Cuffnorm) for each gene using the GenomicRanges R package. Gene essentiality scores for the erythroid K562 cell line were obtained from the CRISPR screen performed by Wang et al. (2015). As the key erythroid transcription factor GATA1 was the most K562-specific essential gene (compared to 3 other chronic myelogenous leukemia cell lines), we determined that the essentiality scores in K562 cells were likely relevant to our primary human erythroid cells. A random forest model was used to determine the percentage of variation in gene expression using Δ mRNA, shLuc mRNA expression, shLuc TE, CDS length, 5′ UTR length, 5′ UTR complexity (Δ G), uAUG presence, and TOP-like motif presence. The random forest was trained on 3,000 genes with measurements for all characteristics and results are reported from the held out set of 618 genes. The R package randomForest was used with the parameters "mtry=3, mtree=200, ntree=501."

Motif Analyses

First, we investigated whether TOP or TOP-like motifs were present within the first 20 nucleotides of the empirical 5′ UTR by matching the strings C(C|U){6} (Thoreen et al., 2012) or (C|U){3}U(C|U){3} (Hsieh et al., 2012). Although we saw an enrichment for TOP-like motifs in RPH-sensitive transcripts, this motif was not present in the majority of transcripts, so we performed a global *de novo* motif analysis of 5′ UTRs (restricted to 30 nucleotides at the 5′ end, 30 nucleotides at the 3′ end, and across the entire 5′ UTR) using Homer with standard options except for "-rna" (http://homer.ucsd.edu/homer/). Next, we took an alternative approach and trained a gapped

k-mer support vector machine (SVM) to try to separate RPH-sensitive transcripts from unchanged transcripts based upon the presence of kmers of length 6, 8, or 10 in the corresponding 5′ UTRs using 5-fold cross validation.

Bone Marrow Biopsy Section Immunohistochemical Staining and Analysis

For each of seven different DBA patients and three normal healthy controls, bone marrow biopsy sections were immunohistochemically stained for GATA1, as previously described (Lee et al., 2017). All sections were stained and imaged together to ensure consistency between samples. Several independent images from each stained sample were segmented and quantified in CellProfiler (Carpenter et al., 2006). In brief, nuclei were segmented by blue intensity and filtered for Hue to retain only brown staining GATA1 positive cells, which we manually confirmed were entirely composed of erythroid cells. We excluded large megakaryocytes by the segmentation procedure. Measurements of intensity and morphological properties were quantified for every cell. Python was used to analyze cellular features, and the Mann-Whitney U non-parametric test was used to estimate the significance of differences observed between DBA and normal cells.

Statistical Analyses

All pairwise comparisons were assessed using an unpaired two-tailed Student's t test, unless otherwise indicated in the main text or in the figure legends. Results were considered significant if the *P value* was < 0.05.

DATA AND SOFTWARE AVAILABILITY

Accession Codes

The accession number for the raw mass spectrometry data reported in this paper is MassIVE: MSV000080283. The accession number for the RNA-seq and ribosome profiling data reported in this paper is GEO: GSE89183.

Cell

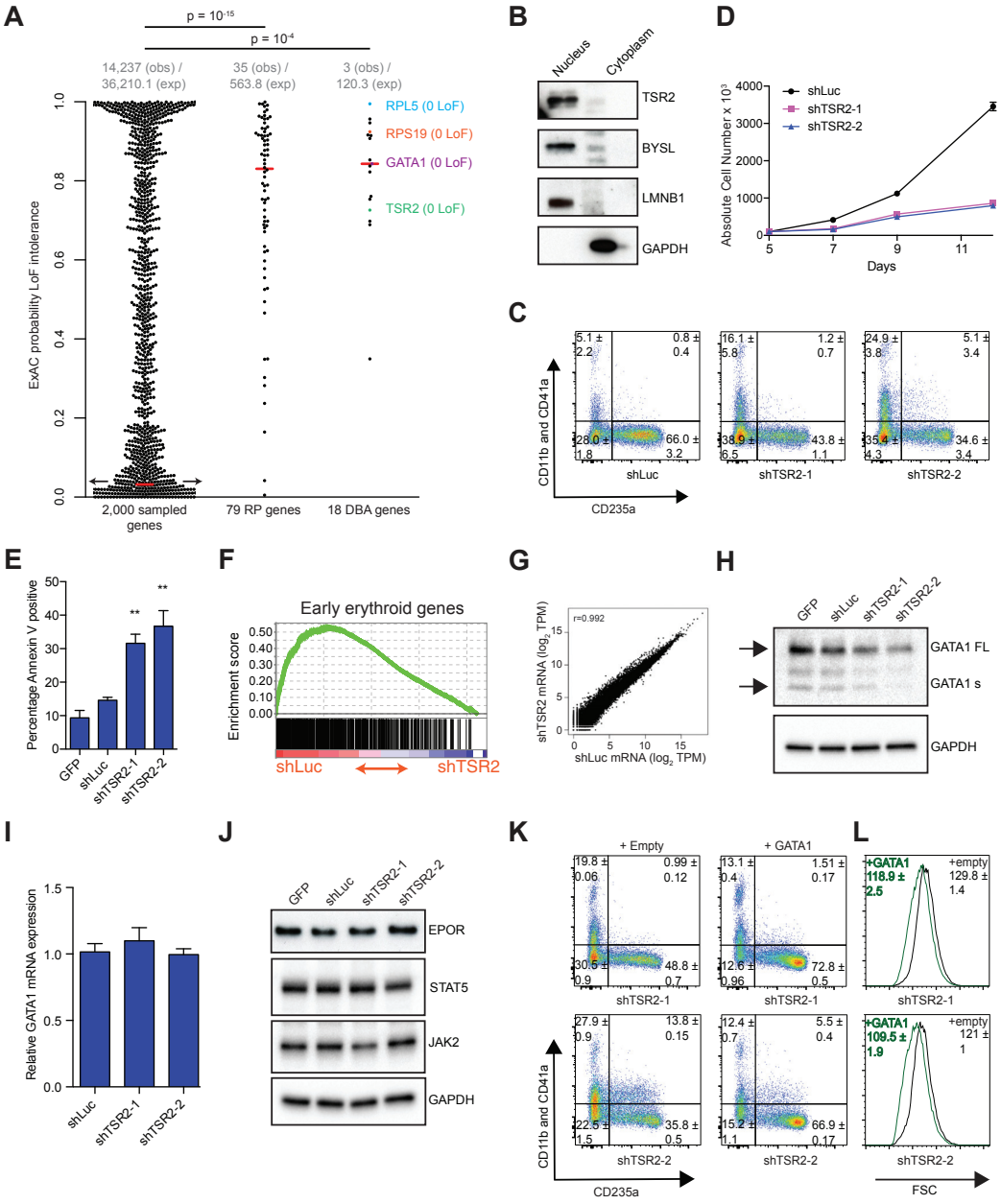


Figure S1. Loss of TSR2 Function Results in a DBA Typical Erythroid Differentiation Defect that Can Be Rescued by GATA1, Related to Figure 1

(A) Similarly to other known DBA genes, TSR2 is significantly de-enriched for loss-of-function (LoF) mutations in 60,706 controls derived from the Exome Aggregation Consortium (ExAC). Note from the mean that the tails of the distribution (intolerant and not intolerant to LoFs) are severely truncated and actually extend much further than can be plotted in comparison to other groups. P values are derived from Mann-Whitney-U tests.

(B) Western blots of nuclear and cytoplasmic human erythroid cell protein lysates for the indicated proteins showing that TSR2 is entirely located in the nucleus.

(C) Representative FACS plots on day 5 after transduction with shTSR2 showing impaired erythroid differentiation of primary human HSPCs *in vitro* and skewing toward non-erythroid lineages. Erythroid cells are marked by CD235a, non-erythroid cells are marked by the expression of CD41a, CD11b, or expression of no markers. Percentages of each subpopulation are shown as the mean ± SD of three independent experiments of cells from three different donors.

(D) Growth curves for primary human HSPCs undergoing erythroid differentiation transduced with shLuc or shTSR2 measured in absolute cell numbers. Shown is the mean ± the SD of three replicates.

(legend continued on next page)

(E) Increased Annexin V staining with TSR2 suppression. Results are shown as the percentage of Annexin V positive cells on day 5 after transduction of primary human HSPCs undergoing erythroid differentiation with shTSR2 or shLuc. Shown is the mean ± the SD of three independent experiments. (**$p \leq 0.01$ using an unpaired two-tailed Student's t test).

(F) Based upon GSEA, cells with TSR2 suppression exhibit a more immature erythroid expression profile (permutation FDR < 0.0001). The enrichment score is plotted in green, and genes are plotted as black lines according to their rank.

(G) Scatterplot of mean gene expression values in shTSR2 and shLuc treated primary human HSPCs undergoing erythroid differentiation on day 5 after transduction.

(H) Western blot detection of GATA1 protein from lysates of human erythroid cells on day 5 after transduction with shTSR2 or shLuc. Arrowheads indicate GATA1 full length and GATA1 short proteins, respectively.

(I) *GATA1* mRNA levels by quantitative RT-PCR (normalized to β-actin) in human erythroid cells on day 5 after transduction with shTSR2 or shLuc. Shown is the mean ± SD of three independent experiments.

(J) Western blot detection of the indicated proteins in human erythroid cell protein lysates on day 5 after transduction with shTSR2 or shLuc showing that the protein levels of other erythroid factors are largely unaffected.

(K) Representative FACS plots of primary human HSPCs undergoing erythroid differentiation on day 5 after transduction with shTSR2 and either with HMD (empty) control or HMD-GATA1 lentiviruses showing that expression of GATA1 rescues the erythroid differentiation defect. Percentages of each subpopulation are shown as the mean ± the SD of three independent replicates.

(L) Representative FACS forward scatter histogram plots (measuring cell size) of cultured primary human HSPCs differentiated toward the erythroid lineage and transduced with shTSR2 and either empty HMD or HMD-GATA1. The forward scatter intensity is shown as mean ± the SD of three independent replicates.

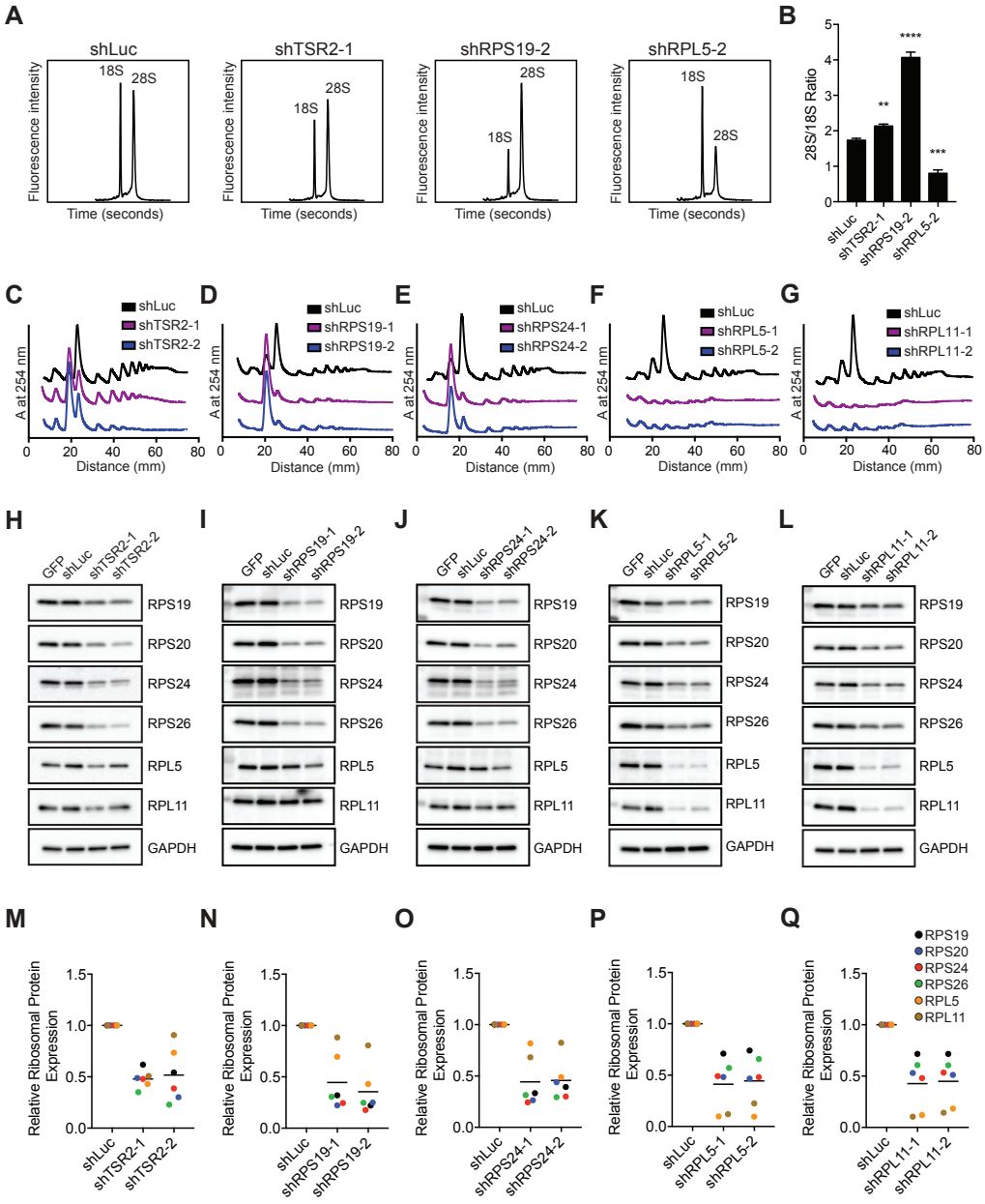


Figure S2. DBA-Associated Molecular Lesions Result in Reduced Ribosome Abundance, Related to Figure 2

(A-B) Bioanalyzer traces of total RNA from human erythroid cells treated with shLuc, shTSR2, shRPS19 (eS19) or shRPL5 (uL18) on day 5 after transduction showing 18S or 28S rRNA processing defects in the respective KDs. Panel A shows one representative of three independent experiments. Panel B shows the mean ± the SD of three independent experiments. (**p ≤ 0.01; ***p ≤ 0.001; ****p ≤ 0.0001 using an unpaired two-tailed Student's t test)

(C-G) Polysome profiles of human erythroid cells on day 5 after transduction showing a reduction of monosomes, polysomes and free amount of the targeted subunit (40S or 60S) with a relative increase of free amount of the non-targeted subunit with indicated DBA-associated molecular lesions. The traces are shown offset from one another on the arbitrary y axis (derived from relative absorbance at 254 nm) for ease of visualizing the data with the x axis showing distance along the sucrose gradient.

(H-L) Western blot detection of the indicated proteins from lysates of human erythroid cells 5 days after transduction with pLKO.GFP, shLuc, shTSR2, shRPS19, shRPS24, shRPL5 or shRPL11 showing the reduction of diverse ribosomal proteins with DBA-associated molecular lesions. Ribosomal proteins of the same subunit as the perturbed protein appear to be more severely affected.

(M-Q) Relative quantification of ribosomal protein band intensities shown in (H-L) using Image Lab.

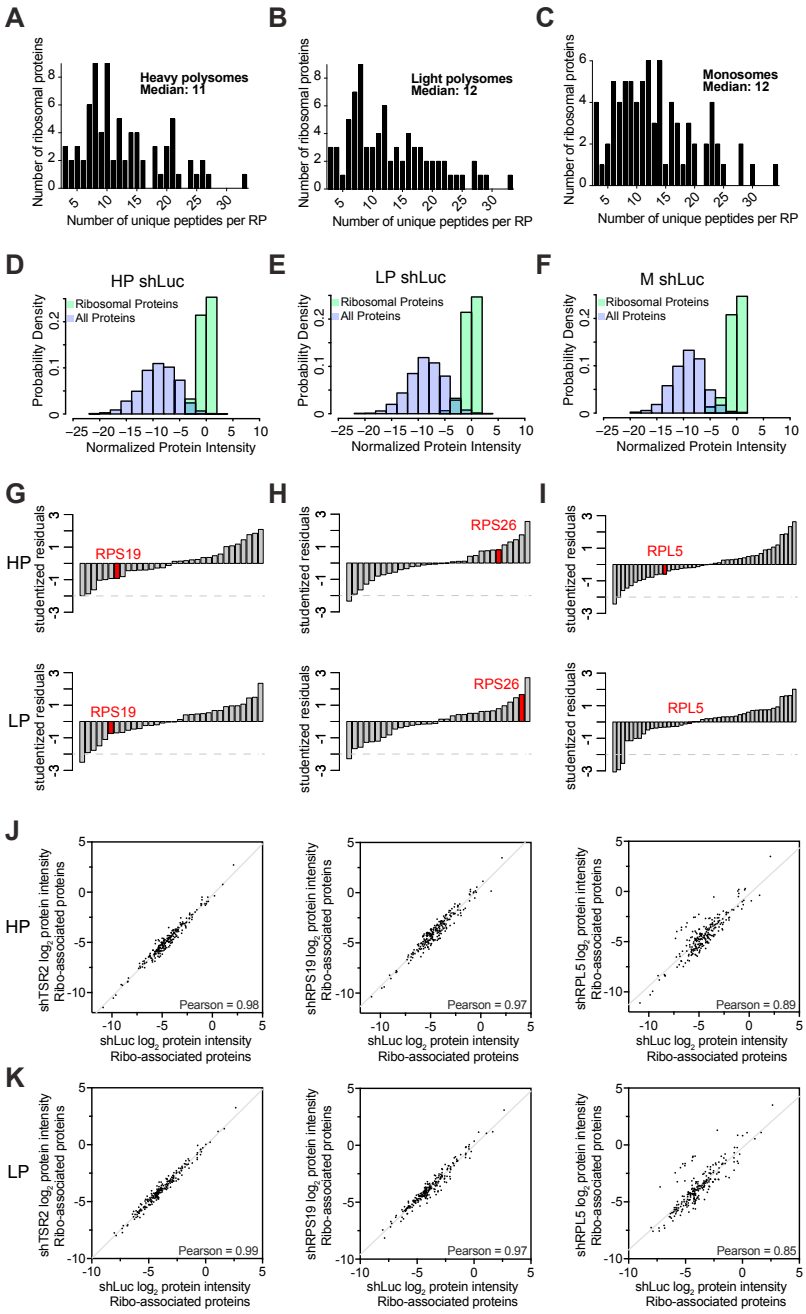


Figure S3. Mass Spectrometry Data Analysis, Related to Figure 3

(A-C) Number of unique peptides quantified per RP in monosomes (M), light polysomes (LP) and heavy polysomes (HP). The median of unique peptides quantified per RP for each fraction is indicated.

(D-F) By sucrose gradient sedimentation, we highly enriched for RPs in HP, LP and M fractions as shown here exemplary for the control samples by probability density, which are representative of all samples.

(G-I) Ordered studentized residual plots are shown for suppression of RPS19, TSR2, or RPL5. Residuals were calculated from the linear fits for the targeted subunit-restricted model (i.e., 40S for shRPS19 and shTSR2 and 60S for shRPL5) shown in Figure 3. The affected RP is highlighted in red for each condition and exhibits no strong deviation in the negative direction from the fit (outliers called at > -3).

(legend continued on next page)

(J-K) The ribosome-associated proteins were analyzed by identifying proteins with similar abundance as RPs in fractions of actively translating polysomes (HP and LP). Log_2 transformed protein intensities from two independent replicates in respective KD condition versus shLuc control in HP and LP fractions showing comparable composition of the ribosome-associated proteins between KD conditions and control. Linear regressions are shown in gray and Pearson correlations are reported. Note that the cluster of proteins that appears to be enriched in the shRPL5 samples consists entirely of eukaryotic translation initiation factors.

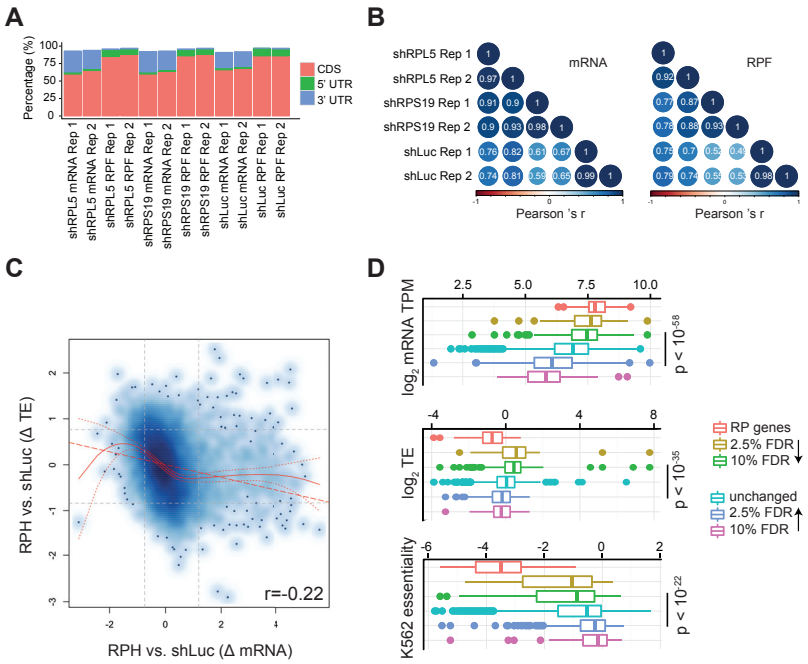


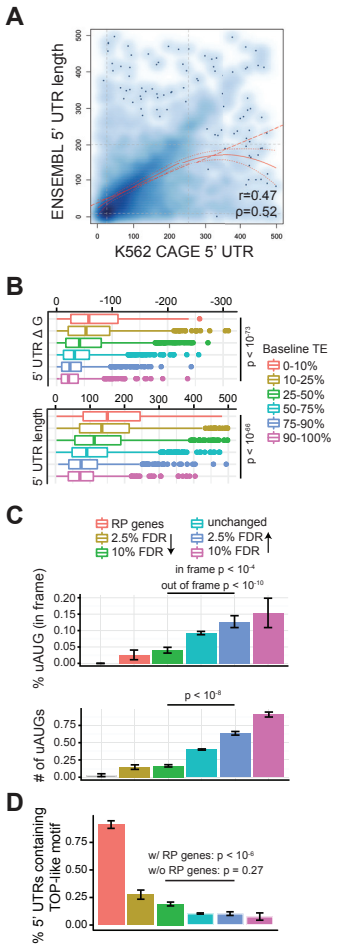
Figure S4. Features of RPH-Sensitive Transcripts, Related to Figure 4

(A) Ribosome protected fragment (RPF) reads are predominately in the CDS and not in the 3′ UTR, whereas mRNAs have a much higher relative percentage of 3′ UTR reads.

(B) Pearson correlations between replicates for $\log_2$ RPF reads (CDS region excluding the first 45 and last 15 nucleotides) and $\log_2$ mRNA reads (entire transcript) are indicated.

(C) Changes in translation efficiency (TE) and mRNA between RPH and shLuc show only limited correlation and are displayed in a scatterplot where color indicates point density. Both local regression (with confidence intervals) and linear fits are shown in red. The Pearson correlation is indicated.

(D) Boxplots for specific features are shown across FDR levels of differential translation. Typical mRNA transcripts per million (TPM) and translational efficiency (TE) are derived from control shLuc cells. Essentiality scores (guide RNA drop out) for K562 erythroid cells were obtained from Wang et al. (2015).



E

5' UTR	L	k	mm	ROCc	PRc
all	6	4	1	0.59	0.19
all	8	6	2	0.58	0.17
all	10	8	3	0.56	0.17
first 30nts	6	4	1	0.51	0.14
first 30nts	8	6	2	0.54	0.15
first 30nts	10	8	3	0.52	0.16
last 30 nts	6	4	1	0.51	0.14
last 30 nts	8	6	2	0.52	0.15
last 30 nts	10	8	3	0.50	0.14

F

entire 5' UTR

Motif	P-value	log P-pvalue	% of Targets	% of Background
GAACCCAC	1e-6	-1.542e+01	4.02%	0.76%
UUCCUCUUCCUG	1e-5	-1.269e+01	8.43%	3.55%
GGUGAGGUUG	1e-5	-1.254e+01	1.61%	0.09%

first 30nts of 5' UTR

Motif	P-value	log P-pvalue	% of Targets	% of Background
UUCCUCUUCCGG	1e-9	-2.237e+01	3.41%	0.25%
UUGGGGGU	1e-8	-2.028e+01	4.62%	0.72%
CCGGGCGAGA	1e-7	-1.710e+01	2.41%	0.15%

last 30nts of 5' UTR

Motif	P-value	log P-pvalue	% of Targets	% of Background
CUGCAGCCAGCC	1e-8	-1.857e+01	1.81%	0.00%
GAAGGUCC	1e-7	-1.718e+01	2.21%	0.10%
GCGGUGUU	1e-6	-1.520e+01	3.01%	0.40%

Figure S5. 5′ UTR Features that Are Associated with Altered Translation Arising from Reduced Ribosome Levels, Related to Figure 5

(A) Annotated 5′ UTR lengths are only moderately correlating with 5′ UTR lengths experimentally determined in erythroid cells by capped analysis of gene expression (CAGE). Both Pearson and Spearman correlations are indicated.

(B) Boxplots for different 5′UTR features are shown across relative baseline TEs in unperturbed primary human HSPCs undergoing erythroid differentiation. P values were determined by an F-test.

(C) Plots for different 5′UTR features are shown across FDR thresholds for differential translation. In-frame and out-of-frame upstream AUGs were determined by string matching in the erythroid 5′ UTR sequences. P values were determined by an F-test.

(D) The percentages of 5′ UTRs containing 5′ terminal oligopyrimidine (TOP)-like motifs within the 20 most 5′ nucleotides are shown across FDR thresholds for differential translation.

(E) Results from the gapped kmer SVM across different regions of the 5′ UTR comparing RPH-sensitive transcripts to all other transcripts. L is the word length, K is the number of informative columns, and mm is the maximum number of mismatches. Area under the receiver operating characteristic curve (ROCc) or area under the precision recall curve (PRc) for each model is shown.

(F) Motifs enriched in the 5′ UTRs of RPH-sensitive transcripts are shown.

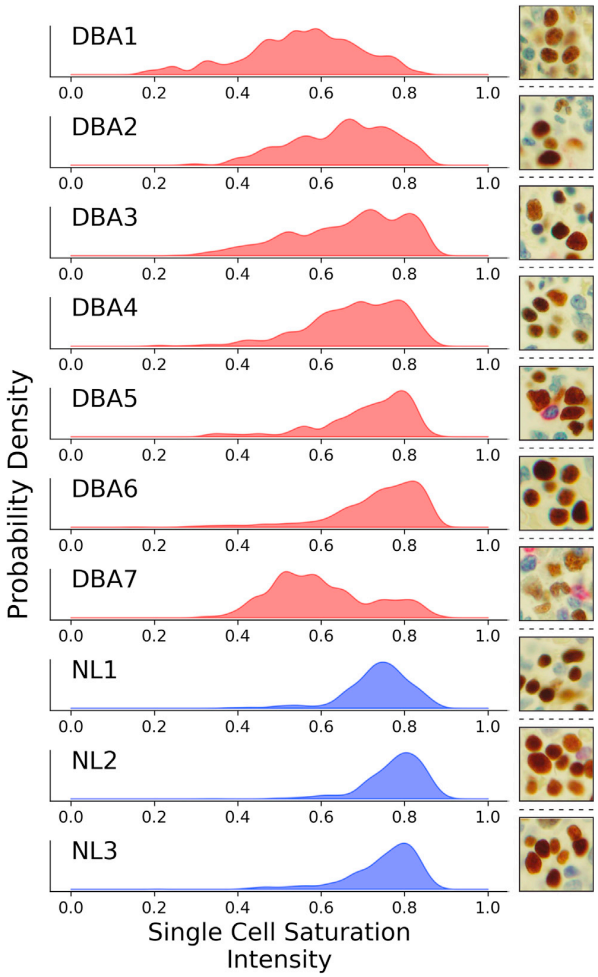


Figure S6. Profiles of GATA1 Immunohistochemical Staining in Bone Marrow Biopsy Sections from DBA Patients and Healthy Individuals, Related to Figure 6

Density plots of saturation intensity at the single cell level for GATA1 across 7 DBA patients and 3 normal individuals. A representative sample image of cells for each patient is provided to the right of each plot.

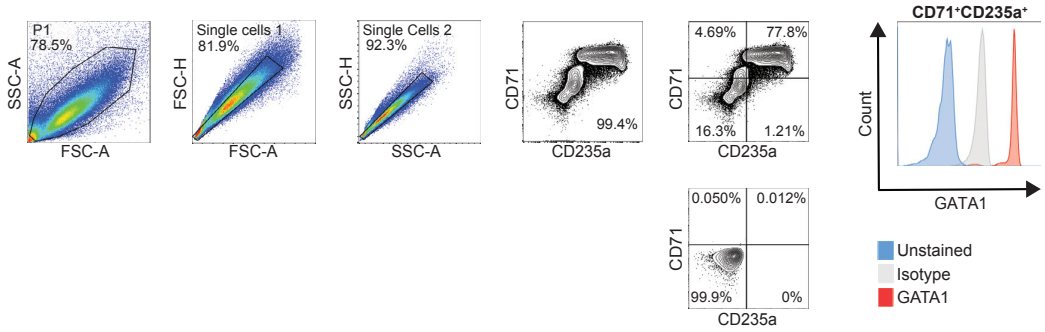


Figure S7. Robust and High-Level GATA1 Protein Expression in Committed Erythroid Cells, Related to Figure 7

Shown is the flow cytometric gating strategy for committed CD235a^{+}CD71^{+} erythroid cells obtained from HSPCs on day 5 of differentiation post-expansion. The cells express high levels of GATA1 protein.

11. Curriculum Vitae/ Lebenslauf

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

12. Complete Publication List

12.1 Publication Presented Within the Scope of This Thesis

1. **Khajuria RK**, Munschauer M*, Ulirsch JC*, Fiorini C*, Ludwig LS*, McFarland SK, Abdulhay NJ, Specht H, Keshishian H, Mani DR, Jovanovic M, Ellis SR, Fulco CP, Engreitz JM, Schütz S, Lian J, Gripp KW, Weinberg OK, Pinkus GS, Gehrke L, Regev A, Lander ES, Gazda HT, Lee WY, Panse VG, Carr SA, Sankaran VG. *Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis.* ***Cell*** *2018 Mar 22;173(1):90-103.e19. doi: 10.1016/j.cell.2018.02.036. Epub 2018 Mar 15*
   (*equal contribution)
   Impact factor (in year of publication): 36.216

12.2 Additional Original Articles and Reviews

2. Kim D*, Park G*, Huuhtanen J, Lundgren S, **Khajuria RK**, Hurtado AM, Muñoz-Calleja C, Cardeñoso L, Gómez-García de Soria V, Chen-Liang TH, Eldfors S, Ellonen P, Hannula S, Kankainen M, Bruck O, Kreutzman A, Salmenniemi U, Lönnberg T, Jerez A, Itälä-Remes M, Myllymäki M, Keränen MAI, Mustjoki S. *Somatic mTOR mutation in clonally expanded T lymphocytes associated with chronic graft versus host disease.* ***Nature Communications*** *2020 May 7;11(1):2246. doi: 10.1038/s41467-020-16115-w.*
   (*equal contribution)
   Impact factor (in year of publication): 11.878

3. Bhere D*, **Khajuria RK***, Hendriks WT, Bandyopadhyay A, Bagci-Onder T, Shah K. *Stem Cells Engineered During Different Stages of Reprogramming*

Reveal Varying Therapeutic Efficacies. ***Stem Cells*** *2018 Feb 16. doi: 10.1002/stem.2805. [Epub ahead of print]*

(*equal contribution)

Impact factor (in year of publication): 5.614

4. Savola P*, Kelkka T*, Rajala HL, Kuuliala A, Kuuliala K, Eldfors S, Ellonen P, Lagström S, Lepistö M, Hannunen T, Andersson EI, **Khajuria RK**, Jaatinen T, Koivuniemi R, Repo H, Saarela J, Porkka K¶, Leirisalo-Repo M¶, Mustjoki S¶. *Somatic mutations in clonally expanded cytotoxic T lymphocytes in patients with newly diagnosed rheumatoid arthritis.* ***Nature Communications*** *2017 Jun 21;8:15869. doi: 10.1038/ncomms15869.*

(*equal contribution; ¶equal contribution)

Impact factor (in year of publication): 12.353

5. Ludwig LS*, **Khajuria RK***, Sankaran VG. *Emerging cellular and gene therapies for congenital anemias.* ***American Journal of Medical Genetics Part C Semin Med Genet*** *2016 Oct 28. doi: 10.1002/ajmg.c.31529.*

(*equal contribution)

Impact factor (in year of publication): 5.600

6. Oh ME, Driever PH, **Khajuria RK**, Rueckriegel SM, Koustenis E, Bruhn H, Thomale UW. *DTI fiber tractography of cerebro-cerebellar pathways and clinical evaluation of ataxia in childhood posterior fossa tumor survivors.* ***Journal of Neuro-Oncology*** *2017 Jan;131(2):267-276. doi: 10.1007/s11060-016-2290-y.*

Impact factor (in year of publication): 3.060

7. Polfus LM*, **Khajuria RK***, Schick UM*, Pankratz N, Pazoki R, Brody JA, Chen MH, Auer PL, Floyd JS, Huang J, Lange L, van Rooij FJ, Gibbs RA, Metcalf G, Muzny D, Veeraraghavan N, Walter K, Chen L, Yanek L, Becker LC, Peloso GM, Wakabayashi A, Kals M, Metspalu A, Esko T, Fox K, Wallace R,

Franceschini N, Matijevic N, Rice KM, Bartz TM, Lyytikäinen LP, Kähönen M, Lehtimäki T, Raitakari OT, Li-Gao R, Mook-Kanamori DO, Lettre G, van Duijn CM, Franco OH, Rich SS, Rivadeneira F, Hofman A, Uitterlinden AG, Wilson JG, Psaty BM, Soranzo N, Dehghan A, Boerwinkle E, Zhang X, Johnson AD, O'Donnell CJ, Johnsen JM, Reiner AP¶, Ganesh SK¶, Sankaran VG¶. *Whole-Exome Sequencing Identifies Loci Associated with Blood Cell Traits and Reveals a Role for Alternative GFI1B Splice Variants in Human Hematopoiesis.* ***American Journal of Human Genetics*** *2016 Aug 4;99(2):481-8. doi: 10.1016/j.ajhg.2016.06.016.*

(*equal contribution; ¶equal contribution)

Impact factor (in year of publication): 9.025

8. **Khajuria RK***, Blankenburg F*, Wuithschick I*, Rueckriegel S, Mansour M, Hernáiz Driever P. *Morphological brain lesions of pediatric cerebellar tumor survivors correlate with inferior neurocognitive function but do not affect health-related quality of life.* ***Child's Nervous System*** *2015 Apr;31(4):569-80. doi: 10.1007/s00381-015-2635-4.*

(*equal contribution)

Impact factor (in year of publication): 1.080

9. Choi SH*, Tamura K*, **Khajuria RK**, Bhere D, Nesterenko I, Lawler J, Shah K. *Antiangiogenic Variant of TSP-1 Targets Tumor Cells in Glioblastomas.* ***Molecular Therapy*** *2015 Feb;23(2):235-43. doi: 10.1038/mt.2014.214.*

(*equal contribution)

Impact factor (in year of publication): 6.938

12.3 Book Chapter

10. **Khajuria RK**, Gross BA, Du R. "*Image-Guided Open Cerebrovascular Surgery"* in *Image-Guided Neurosurgery*. 1st ed. Ed. Alexandra J. Golby. Boston: Academic Press, 2015. 277-96

13. Acknowledgements

I am very grateful for having had the opportunity to work under the guidance of Vijay G. Sankaran supported by Petra Knaus and Stefan Mundlos in an exciting, fast paced, cutting-edge and probably one of the best possible research environments in the world.

I am particularly grateful for Vijay's mentorship, I received from him the right balance between the freedom to develop myself scientifically and also guidance where and when necessary. I am also thankful for his support and understanding for matters outside the lab. I thank Petra, Stefan and Sabine Bartosch very much for supporting my aspiration to become part of the MD/Ph.D. program at the BSRT.

I would also like to thank my colleagues, labmates and collaborators, many of which became good friends of mine, for this incredible experience I had while working on this project and beyond. This all would not have been possible without your help, support and also the fun that we had together: Leif S. Ludwig, Jacob C. Ulirsch, Mathias Munschauer, Claudia Fiorini, Marko Jovanovic, Harrison Specht, Sean K. McFarland, Nour J. Abdulhay, Jesse M. Engreitz, Charles P. Fulco, D.R. Mani, Hasmik Keshishian, John Lian, Satish Nandakumar, Anindita Basak, Baris Boyraz, Arshed Al-Obeidi, Akshay D'Cruz, Rockwell Anyoha, Aryeh Solomon, Liz Perez, Klara Sirokman, Venkat Magupalli, Roberto Negro and Marko Knoll.

Last but not least I would like to thank the people, who made this all possible, my family. This work is dedicated to you all. I thank my wife, who sacrificed any own aspirations and is always backing me so I can follow and live my dream, I love you Maija. Sophie and Sami, you are the meaning of my life, your health and happiness is all that matters to me for the rest of my life and lets me cope with any hurdles that I may face in and outside the academic life. My parents Erika Ingrid Khajuria and Sushil Kumar Khajuria as well as my brothers Chandan Kumar and Ranjan Kumar made me the person, that I am today. I thank you for your life-long support and for always being there for me, I love you. I always wanted to make you all proud and hope that I achieved that at least a little bit with this work.