Abstract

During red blood cell (RBC) and platelet formation in mammals, translation proceeds in the absence of new mRNA synthesis. The mechanisms of translational control operating during these phases of non-nucleated blood cell terminal differentiation are incompletely understood. Here, we comprehensively profiled translation by maturing RBCs and platelets using ribosome profiling, expanding the role that should be anticipated for translation and its control in these cell types. Next, using in vitro models of erythroid and megakaryoid differentiation, we present evidence supporting a differentiation-dependent switch in the machinery responsible for dissociating (“recycling”) ribosomes from mRNAs after translation termination. This switch involves the loss of ABCE1-mediated ribosome recycling, which is compensated by upregulation of PELO/HBS1L. We demonstrate, in the early phases of differentiation of K562 cells, that PELO/HBS1L induction promotes hemoglobin synthesis at the translational level by increasing the availability ribosomes. In the latest phase of differentiation, we show that there is a nearly complete loss of ribosome recycling factors leading the dramatic increase in occupancy of 3’UTRs by ribosomes.
Nicholas T. Ingolia (Sponsor and Reader)
Assistant Professor
Department of Molecular and Cell Biology
University of California, Berkeley

Rachel Green (Sponsor and Reader)
Professor
Department of Molecular Biology and Genetics
Johns Hopkins University School of Medicine

Geraldine Seydoux (Reader)
Professor
Department of Molecular Biology and Genetics
Johns Hopkins University School of Medicine
Acknowledgments

I would like to thank Nick Ingolia and Rachel Green for their guidance and support. The opportunity to work with two excellent mentors has been humbling and inspiring. The complementarity of their unique approaches towards science has benefited my development as a scientist enormously. Nick and Rachel deserve outstanding credit for this. Their combined cleverness, creativity, and systematic thinking are traits which are naturally cultivated in their students by working with them, and traits which I hope to further develop in my future endeavors. I would like to thank Geraldine Seydoux for always providing an exacting outside perspective – during thesis meetings or just stopping in the hallway. Her curiosity about science is contagious. I would also like to thank Mary Armanios, Sarah Woodson, and Gregg Semenza and for their input and helpful advice. I would like to thank all of the Ingolia lab members, but especially Anna McGeachy and Mike Harris, as well as colleagues from the Carnegie Institution – Gaëlle Talhouarne, Fred Tan, Valeriya Gayinskaya, Allan Spradling. I would like to thank all of the Green lab members, especially Jamie Wangen who helped out with key experiments. I would like to thank my parents, Cheryl and Bill Goodgine & Charles and Gerda Mills, without whom the opportunity to do this work would not have been possible. I am greatly indebted to their love and selflessness. Finally, my wife Huan – your support and encouragement has been inestimable. But more than that, the strength and tenacity you have channeled in your own pursuits has throughout graduate school been, for me, a source of unfolding inspiration, a height to aspire towards.
List of Figures

Chapter I

Figure 1. Red blood cell and Megakaryocyte differentiation

Figure 2. Overview of Translation

Chapter II

Figure 1. Hematopoietic cell lineage map

Figure 2. Platelets and reticulocytes accumulate 3’UTR ribosomes

Figure 3. 3’UTR ribosomes accumulate during in vitro differentiation of Meg01 and K562 cells

Figure 4. 3’UTR ribosomes are not translating a polypeptide

Figure 5. Ribosomes in the 3’UTR are non-translating post-termination ribosomes

Figure S1. CD42p expression in plateletpheresis samples

Figure S2. CD41 expression by Meg01-derived platelet-like particles (PLPs)

Figure S3. RPF scatter plots of primary platelets and CD71-selected reticulocytes

Figure S4. Proteins previously reported to be synthesized in platelets

Figure S5. Gene ontology of highly expressed platelet mRNAs

Figure S6. Correlation among global gene expression measurements in platelets

Figure S7. Translational efficiency (TE) of platelet mRNAs

Figure S8. Ribosomes are depleted in platelets

Figure S9. Platelet activation by thrombin or ADP

Figure S10. Translational changes following platelet activation
Figure S11. RPF scatter plots of activated and resting platelets

Figure S12. Many platelet mRNAs are enriched relative to megakaryocytes

Figure S13. De novo motif searches of platelet-enriched mRNAs

Figure S14. Other identified RNA motifs and interacting-RBPs

Figure S15. FMRP is expressed by Meg01 cells

Figure S16. FMRP is important for expression of a subset of PLP-enriched mRNAs

Figure S17. FMR1-null mice have abnormal megakaryocyte differentiation

Chapter III

Figure 1. Ribosome rescue/recycling factors are regulated during differentiation

Figure 2. Translational arrest by PLPs

Figure 3. Restoration of PELO/HBS1L releases 3’UTR ribosomes in PLPs

Figure 4. PLP gene models

Figure 5. Restoration of PELO/HBS1L in differentiated K562 cells

Figure 6. Clearance of 3’UTR ribosomes is associated with mRNA decay

Figure 7. Translation, not RNA abundance, drives hemoglobin accumulation in K562 cells

Figure 8. ABCE1 silencing leads to ribosome recycling failure and is rescued by PELO/HBS1L over expression

Figure 9. PELO/HBS1L induction promotes fetal hemoglobin synthesis by increasing free ribosomes

Figure 10. Model – A differentiation-dependent switch in ribosome recycling factors favor hemoglobin translation
Figure S1. Purification of transgenic PLPs
Figure S2. Evidence for hyperphosphorylated forms of HBS1L
Figure S3. PELO/HBS1L overexpression interferes with normal XBPI translation

Appendix I

Figure 1. Cap-independent translation assay using a covalently circular RNA
Figure 2. Circular RNA IRES reporter for library screening
Figure 3. High throughput IRES profiling strategy
Figure 4. \textit{in vitro} ribozyme-mediated cyclization
Figure 5. Nascent chain-mediated translational pausing in IRES reporter
Figure 6. Functional validation using the VEGFA IRES \textit{in vitro} and in HEK293
Figure 7. Ribozyme-mediated IRESseq library cyclization and expression in HEK293 cells
Figure 8. IRESseq enrichment for specific library sequences
Figure 9. Translational control during hypoxia
Figure S1. IRESseq input library complexity
Figure S2. IRESseq spike-in control sequences
Figure S3. Translational changes during hypoxia and reoxygenation
Chapter I: Introduction

Introduction

The central dogma of molecular biology describes the flow of genetic information in all cells from its storage form of DNA, to short-lived mRNA transcripts, and finally its translation into proteins by the ribosome. In multicellular organisms, this process undergoes modulation during the diversification of cell types\textsuperscript{1,2}. During differentiation, changes in the gene expression ultimately lead to functional specialization\textsuperscript{3}. But what happens when functional specialization requires eliminating parts of the cell’s gene expression machinery? In mammalian tissues, terminal differentiation is often accompanied with loss of specific organelles, for example, the nuclei of lens fiber cells\textsuperscript{4}, keratinocytes\textsuperscript{5}, platelets and mammalian red blood cells (RBCs) (Fig. 1a)\textsuperscript{6,7}. The unique ultrastructural changes that occur during erythroid and megakaryoid blood cell differentiation are reviewed here with an emphasis, where possible, on how gene expression is regulated during these processes.

Mammalian RBCs and platelets are non-nucleated

In mammals, the smallest mammalian blood vessels (~3.5 µm) have diameters half the size of a red blood cell (~7 µm)\textsuperscript{8,9}. Red blood cells are the oxygen-carrying cells of blood and must be physically deformed to pass through these vessels – a fact which is believed, through evolution, to have resulted in mammals having non-nucleated red blood cells\textsuperscript{7}. Platelets are non-nucleated cell fragments that react to sites of damaged blood vessels, forming blood clots, and arresting the blood loss to the extracellular space.
Damage to the blood vessel wall leads to changes in local shear forces that are sensed by circulating platelets which then become activated and anneal directly to the exposed subendothelium\textsuperscript{10-12}. Platelet aggregation is potentiated by ADP release, and reduces blood flow to the area of the injured vessel\textsuperscript{13}, limiting (and ultimately stopping) blood loss. In all other vertebrates, cells which serve a similar function to platelets are nucleated and also function as immune cells\textsuperscript{14}. Thus, the ancestral function of platelets as dual immune and blood clotting cells provides rationale for their secondary function in inflammation and immunity\textsuperscript{15,16}, despite most of these functions in mammals being relegated to white blood cells. However, the reason for anucleate platelets in mammals is unknown.

**Transcriptional and translational arrest during RBC maturation**

Nuclear elimination by RBCs and platelets occurs by distinct, unrelated mechanisms. In adult mammals, nucleated erythroid precursor cells (erythroblasts) are the first to begin to synthesize hemoglobin (Hb) ([Fig. 1a](#)). During early erythroblast differentiation, precursors are mitotically active and Hb synthesis occurs primarily during the G\textsubscript{1} and early S phase of the cell cycle\textsuperscript{17}. In late erythroblasts, chromatin condensation begins and cells undergo a period of prolific rRNA transcription and ribosome assembly leading to an increase in total cellular RNA content\textsuperscript{7,18,19}. As the last mitotic cycle is complete, late erythroblasts exit the cell cycle and Hb accumulation intensifies, reflecting the nearly complete shift by these cells towards translating solely Hb mRNA\textsuperscript{20}. Importantly, transcription of RNA ceases in late precursor cells prior to enucleation\textsuperscript{17,19}. The reticulocyte stage of erythroid maturation is, thus, defined by the extrusion of the cell
nucleus (or the “pyrenocyte”\textsuperscript{21}). In these cells, mRNAs are stably associated with highly abundant polysomes\textsuperscript{19} and Hb mRNA comprises >70\% of the transcriptome\textsuperscript{22}.

Reticulocyte protein synthesis occurs for \textasciitilde24-36 hours in circulation, during which time one third of the final amount of hemoglobin protein present in the mature red blood cell is be synthesized. After synthesis, the hemoglobin protein is stable for the entire 120 lifespan of the mature red blood cell.

In newly released peripheral blood reticulocytes, \textasciitilde80-90\% of all ribosomes are found in polsosome fractions consisting of 2-6 ribosomes\textsuperscript{23-26}. During reticulocyte maturation, the fraction of total ribosomes found in polsosome clusters may vary (some studies report an increase\textsuperscript{27,28} while others report a decrease\textsuperscript{19,23,26}). However, because overall ribosome abundance decreases \textasciitilde10-20 fold over 24 hours, and the disaggregation of polsosomes occurs slowly if at all, it is likely that reticulocyte polsosomes are highly stable structures which – when released – are rapidly degraded. Interestingly, isotopic incorporation studies of isolated reticulocyte polsosomes suggests that a population of inactive polsosomes accumulate during reticulocyte maturation that are not engaged in protein synthesis\textsuperscript{28-30}. More recent work \textit{in vitro} has suggested this ribosome population could reflect translationally silenced mRNPs that lack polyA-binding protein (PABP)\textsuperscript{31}, in whose absence (in yeast) a similar phenotype results\textsuperscript{32}. It has been suggested that the accumulation of inactive polsosomes accounts for translational arrest by reticulocytes rather than disaggregation of polsosomes themselves\textsuperscript{28-30}.

\textbf{The formation of platelets}

Platelets, unlike their erythroid counterparts, do not arise from nucleated precursor cells
that eliminate their nuclei through gradual chromatin condensation. Instead, platelets originate from the cytoplasmic fragmentation of rare\(^1\) polyploid precursor cells called megakaryocytes (MKs)\(^ {33}\) (Fig. 1a). Committed MK precursors undergo between 2 and 6 rounds of endomitosis leading, on average, to immature MKs with a ploidy of 16N (but can be as high as 128N)\(^ {34}\). RNA and DNA fluorescence \textit{in situ} hybridization (FISH) analyses have demonstrated that all alleles of MKs are active\(^ {35}\), and gene expression profiling of MKs sorted by ploidy level revealed remarkably little differential gene expression\(^ {36}\). This indicates that MK polyploidization most likely acts as a form of functional gene amplification required for the rapid increase in transcription and translation needed to accompany the massive (> 10 fold) increase in cell size by mature (platelet-producing) MKs\(^ {14,35,36}\). Notably, this outcome differs from the effects of polyploidization in other systems, where gene dosage is apparently tightly controlled after polyploidization by selective silencing of alleles\(^ {37}\). The broad increase in expression of all genes by differentiating MKs stands as a point of contrast with the developing erythroblasts/reticulocytes in which the gene expression machinery is necessarily dominated by the synthesis of a single gene product (i.e. hemoglobin).

MKs next exit the cell cycle and rapidly acquire the ultrastructural features of platelet-producing cells. This involves the synthesis and transport of thousands of alpha granule (and other) platelet proteins as well as the formation of a highly intricate demarcation membrane system (DMS) that defines future nascent platelet “zones”\(^ {38,39}\). The detailed events which lead to trafficking and sorting of MK components to nascent platelets are only beginning to be understood\(^ {40}\). However, when platelets are released,

\(^1\) MKs comprise only 0.02-0.05% of bone marrow cells, but together release \(1 \times 10^{11}\) platelets per day. Each MK is thought to release \(~3000\) platelets. See Chernoff \textit{et al.}, (1980).
they contain ribosomes\textsuperscript{41-44}, and like reticulocytes, synthesize protein\textsuperscript{45-53} using a MK-derived pool of mRNAs\textsuperscript{54-56}. Unlike the established role for reticulocyte protein synthesis to synthesize Hb molecules that will be long-lived during the 120 day RBC lifespan, a functional role for platelet protein synthesis during their short 5-9 day lifespan is unknown. Although the effects of translation inhibitors in \textit{ex vivo} platelet function assays has generally supported a role for translation in platelet aggregation\textsuperscript{57-59}, deciphering the role of translation by platelets in more complex processes such as innate immunity\textsuperscript{16}, atherothrombosis\textsuperscript{15}, and cancer metastasis\textsuperscript{60,61} will prove important future research goals.

**The translation cycle and its mechanism**

The translation of proteins in eukaryotes begins with the formation of a preinitiation complex consisting of the 40S ribosome and a ternary complex consisting of the eukaryotic initiation factor eIF2 bound to GTP (eIF2-GTP) and the initiator tRNA\textsubscript{i}Met\textsuperscript{62}. This reaction is assisted by eukaryotic initiation factors 1, 1A, 3, and 5 and leads to the recruitment of the scanning ribosomal complex by the mRNA cap binding complex eIF4A, 4G, and 4E (Fig. 2a). The start codon (usually the first AUG) is identified by base pairing with the Met-tRNA\textsubscript{i}Met in the ribosomal P site and facilitated by a specific nucleotide context (Kozak Sequence)\textsuperscript{63}. 60S subunit joining leads to the formation of an elongation-competent 80S ribosome and amino acid incorporation (Fig. 2b). Translation termination occurs when the ribosome encounters a stop codon (UAA, UAG, UGA) in the ribosomal A site. Eukaryotic release factor 1 catalyzes the hydrolysis of the peptidyl-tRNA, releasing the nascent polypeptide chain from the ribosome (Fig. 1c). It is assisted in this role by eRF3 which shuttles eRF1 to the ribosome. The unusual ABC transporter
family member, ABCE1 (which lacks transmembrane segments and thus transporter activity) stimulates nascent peptide hydrolysis by eRF1 (Fig. 1c) and is required for subunit dissociation from the mRNA after release of the nascent protein – this process is referred to as ribosome recycling because it returns the ribosomal subunits to the cytoplasmic pool of ribosomes available for additional round of translation. Additional eukaryotic release factor paralogues have been identified that are involved in termination-like ribosome release reactions but that operate independently of a stop codon. These processes are collectively termed “ribosome rescue” because they occur in order to remove stalled ribosomes from an mRNA and thus return them to the available pool of ribosomes for additional rounds of translation. Ribosome rescue involves PELO (Dom34 in yeast) and HBS1L (Hbs1 in yeast) which are paralogues of eRF1 and eRF3, respectively (Fig. 1d). Like eRF1/3, PELO/HBS1L bind to the vacant A site of a translationally stalled ribosome and, through a similar interaction with ABCE1, act to release the stalled subunits form the mRNA.

**General and selective translational control mechanisms**

Maturing red blood cells at the reticulocyte stage of development are among the most well-studied cell types in all of biology. The ability to reconstitute protein synthesis in reticulocyte extracts following the removal of endogenous mRNAs and the addition of in vitro prepared RNA templates defined by the user led to a veritable explosion in the understanding of translation mechanism and its control. The study of reticulocytes thus has contributed significantly toward defining the major paradigms of translational control. Global mechanisms involve changes in the translation of all mRNAs. These
mechanisms often act by increasing or decreasing the rate of translation initiation. For example, heme, distinct from its role as a prosthetic group of hemoglobin, is a translational regulator that stimulates protein synthesis. In the absence of heme, binding of methionyl-tRNA$_{\text{Met}}$ to 40S subunits is blocked. The mechanism of repression involves activation of a heme controlled repressor (HCR) which functions as an eIF2$\alpha$ kinase. Reversible phosphorylation of eIF2$\alpha$ on Ser-51 is a major translational control pathway integrating diverse cellular stress signals in all eukaryotes leading to the global reduction of translation initiation. Other eIF2 kinases include a RNA-dependent protein kinase (PKR), PKR-like ER kinase (PERK), and GCN2 (general control non-depressible-2). Phosphorylation of eIF2$\alpha$ inhibits translation, not by directly blocking eIF2 function, but instead by interfering with the guanine nucleotide exchange factor (GEF) eIF2B-mediated recycling of eIF2-GDP which must occur prior to Met-tRNA$_{\text{Met}}$ binding. eIF2 expression levels are substantially higher than eIF2B and, thus, a small amount of phosphorylated eIF2[\alphaP] becomes an effective competitive inhibitor of eIF2B. eIF2 kinases are broadly conserved and responsible for mediating stress-induced translational repression in response to heme-deprivation (HCR), viral infection (PKR), unfolded protein stress (PERK), and serum/amino acid starvation (GCN2).

Selective mRNA translational control mechanisms have been uncovered more slowly owing to their (often) more complex underlying mechanisms and it is becoming increasingly appreciated that general translational control mechanisms often coexist with mRNA-specific effects. Global translational repression following activation of eIF2$\alpha$ kinases, for example, selectively stimulates the translation of specific mRNAs such as
GCN4, ATF4, and C/EBP isoforms. The mechanism through which increased levels of eIF2[αP] stimulates translation this subset of mRNAs involves the translation of several upstream open reading frames (uORFs) that are translated by a “leaky scanning” mechanism. It is thought that ribosomes translating a short uORF are capable, after termination, of resuming scanning (rather than dissociating from the mRNA) – and that the subsequent “distance” (measured in number of nucleotides) such ribosomes are able to scan thereafter before reinitiating depends crucially on the availability of Met-tRNA_{Met}-eIF2-GTP ternary complex. Such reinitiating ribosomes are unable to initiate without first reacquiring a tRNA_{Met}-eIF2-GTP ternary complex. Thus in the case(s) of GCN4, ATF4, and C/EBP, mRNA-specific translation is stimulated by low levels of the tRNA_{Met}-eIF2-GTP ternary complex because it permits longer distance scanning of reinitiating ribosomes, thus bypassing unproductive uORF initiation, and increases the probability of a scanning ribosome reacquiring a ternary complex nearby to the “correct” AUG to make the functional protein. Despite clear evidence that uORFs “sensitize” mRNAs to the availability of eIF2[αP], the mechanisms governing whether terminating ribosomes are recycled, or resume, scanning are unknown.

A common theme in translational control has thus emerged that selective translational regulation often involves regulation by general translation factors or their paralogues. Along with the availability of eIF2-GTP, cellular levels of the cap binding complex (eIF4F) represents one of the best understood translational control mechanisms. eIF4G is a scaffold protein that interacts with the cap binding protein eIF4E through a conserved domain found in multiple 4E-Binding Proteins (4E-BPs). The phosphorylation state of 4E-BPs regulate their ability to interrupt the eIF4G/4E interaction required for
43S binding to the mRNA\textsuperscript{84}. The activation of 4E-BPs (by dephosphorylation) leads to global repression of cap-dependent translation by interfering with eIF4G/4E binding. Specific mRNAs which lack strict dependence on the cap binding complex eIF4F, such as those containing an internal ribosome entry site (IRES) continue to be translated under these circumstances\textsuperscript{85}. While many cellular IRESes are viewed as having only weak activity in promoting initiation under normal conditions, loss of eIF4F (which can be cleaved by viral proteases\textsuperscript{86-89} or apoptosis factors\textsuperscript{90}) likely favors IRES-mediated translation by increasing the availability of translation initiation components (See Appendix I). Not surprisingly, diverse environmental signals regulate the eIF4G/4E interaction during cancer progression\textsuperscript{91, 92}, viral infection\textsuperscript{86}, and development\textsuperscript{93}.

**mRNA-specific translational control mechanisms**

Unlike general translational control mechanisms described above, *bona fide* mRNA-specific translational control mechanisms are remarkably diverse and involve both specific *trans*-acting factors (e.g. RNA binding proteins\textsuperscript{94, 95}, microRNAs) and regulatory sequences/structures present in the 5’- or 3’-UTR, coding sequence, or nascent polypeptide chain. mRNA-specific translation control mechanisms have been uncovered which act at all steps of the translation cycle (initiation, elongation, termination) except ribosome recycling insofar as the termination/reinitiation of uORFs is not taken to reflect regulation of ribosome release (though this could in principle be the case).

Messenger RNA-specific regulation of initiation can occur by blocking 43S recruitment to specific mRNAs, for example, by iron regulatory proteins (IRPs) in mammals\textsuperscript{96-98}, the *Drosophila* Sex-lethal protein (Sxl)\textsuperscript{99}, or the oocyte protein Maskin in
vertebrates\textsuperscript{100, 101}. In reticulocytes, translation of erythroid lipoxygenase is blocked at the 60S joining stage of initiation by HNRNPs K and E\textsuperscript{72}. In a highly unusual mechanism, the 60S ribosomal protein L13a dissociates from the large subunit in interferon-treated cells and represses translation by interacting with the ceruloplasmin mRNA in the 3´UTR where it prevents eIF4G-dependent recruitment of eIF3\textsuperscript{102, 103}. MicroRNAs, which act to both repress translation and stimulate mRNA decay reduce translation of target mRNAs at a yet unidentified step during translation initiation\textsuperscript{104-109}. Translational control targeting the elongation step of protein synthesis often involves ribosome pausing due to specific regulatory nascent peptide sequences such as those containing polyproline motifs, which in bacteria are relieved by EF-P\textsuperscript{110-112} and in mammalian cells by eIF5A\textsuperscript{113}. Other stalling peptides have been shown to mediate mRNA cellular localization, for example the C-terminus of XBP1 in mammals, which directs ribosome pausing and subsequent trafficking of the entire ribosome-nascent chain complex to the ER membrane before translation resumes\textsuperscript{114}. Translation termination may also be regulated during mRNA-specific translational control. Most well known, selenoproteins are synthesized by a mechanism that involves a recoding event of UGA stop codons that involves both a well-conserved \textit{cis}-acting stem loop structure (the sec-insertion sequence; or SECIS) in the 3´UTR of mRNAs encoding selenoproteins as well as \textit{trans}-acting nonsense suppressor tRNAs and SECIS binding proteins\textsuperscript{115, 116}. The actual process through which recoding occurs, however, is not fully understood. Recently, the vascular endothelial growth factor A (VEGFA) protein was shown to be regulated by translational read through of the normal stop codon\textsuperscript{117}. In this specific case, the resulting proteins produced by either the
predicted full length open reading frame, or the read through product had opposite biological effects.

**Conclusions**

Terminal differentiation is an irreversible switch in cellular behavior and involves the acquisition of specific cellular functions while dispensing with others. In developing RBCs and platelets, translation proceeds in the absence of transcription, and (at least for RBCs), major portions of available cellular resources are devoted to the synthesis of a proteins. The subject of this thesis involves experiments undertaken in an attempt to understand how translation is regulated in cell types in which transcription is arrested. Terminal differentiation of erythroid and megakaryoid blood cell lineages, naturally, are excellent human model systems for this purpose.
References


6. Gulliver, G. *Observations on the sizes and shapes of the red corpuscles of the blood of vertebrates, with drawings of them to a uniform scale, and extended and revised tables of measurements*, Proceedings of the zoological society of London, 1875).


Figure Legends

Figure 1. **Red blood cell and megakaryocyte differentiation** | **a. Upper Lineage.** Stages of erythropoiesis originating from the hematopoietic stem cell (HSC), proceeding through the erythroblast stages, and leading to the formation of CD71+ enucleated reticulocytes (with the extrusion of the pyrenocyte). Reticulocyte maturation completes in circulation resulting in the formation of mature CD71- biconcave red blood cells (RBCs). Mature RBCs circulate for ~120 days. **Lower Lineage.** Megakaryoblasts form from a common myeloid progenitor, and undergo a series of endomitotic cycles leading to the formation of platelet-producing megakaryocytes. Platelets are then released into the blood stream where they circulate for 5-10 days. **b.** Hematopoietic lineages including white blood cell lineages. Lineages highlighted in blue give rise to anucleate forms during terminal differentiation, while those highlighted in pink give rise to nucleated forms. **c.** Mammalian hemoglobin loci on chromosome 16 and chromosome 11. Globin genes are colored in blue and pseudogenes (psi) are colored in gray. Globin chains making up the various hemoglobin tetramers expressed during development are shown (Embryonic, Fetal, and Adult).

Figure 2. **Overview of translation** | **a.** During translation initiation, 40S small subunit ribosomes are bound by cytoplasmic ternary complex (TC), consisting of eIF2-GTP-Met-tRNA$_i^{Met}$. The 40S ribosome binds to the ternary complex under the guidance of the eukaryotic initiation factors eIF1, 1A, 3, and 5, forming a 43S preinitiation complex. The 43S complex is recruited to mRNA by the cap binding complex eIF4F (consisting of...
eIF4A, 4G, and 4E), leading to bind of the 43S complex to the mRNA. This complex then scans down the mRNA and selects the initiation codon. 60S subunit joining results in the formation of elongation competent ribosome complexes. b. Elongation results in the extension of the nascent polypeptide by successive peptidyl transfer reactions. c. Near the end of translation, a stop codon (UAA, UAG, UGA) is encountered in the ribosomal A site and is recognized by eukaryotic release factors (eRFs) 1 and 3. eRF1 catalyzes peptidyl-tRNA hydrolysis and is stimulated by the atypical ABC type ATPase ABCE1, resulting in the release of the completed polypeptide. After termination, ribosome subunits must be dissociated from the mRNA in a process termed “recycling”. Ribosome recycling is primarily performed by ABCE1. d. Stalled ribosomes must also be released by termination-like processes. eRF1/3 homologues PELO/HBS1L recognize stalled ribosomes in a codon-independent manner and stimulate their dissociation (“rescue”) from mRNA along with ABCE1 which participates in ribosome recycling and ribosome rescue processes.
Fig. 1

(a) Bone Marrow

Bone Marrow

Myeloid

Myeloid

HSC

BFU-E

CFU-E

BFU-MK

CFU-MK

pro-megakaryoblast

megakaryoblast (4N)

Megakaryocyte

(4N to 64N)

Platelet-producing megakaryocyte

Blood vessels

Reticulocyte

RBCs

Platelets

(b) Platelets

Megakaryocyte

Platelets

Erythrocytes

Common Myeloid

Neutrophil

Monocyte

Dendritic cell

Mast Cell

Osteoclast

Common Lymphoid

T-cell

B-cell

NK-cell

(c) globin gene

\( \zeta \)

\( \psi \zeta \)

\( \psi a2 \)

\( \psi A1 \)

\( a2 \)

\( a1 \)

\( \theta 1 \)

\( \zeta F2 \)

\( \zeta Y2 \)

\( \alpha F2 \)

\( \alpha Y2 \)

\( \alpha S2 \)

\( \alpha A2 \)

Hb Gower 1

Hb Portland

Hb Gower 2

Hb F

Hb A

Embryo

Fetus

Adult

chr16

\( \epsilon \)

\( \gamma Y \)

\( \gamma S \)

\( \alpha Y \)

\( \alpha S \)

chr11

\( \delta \)

\( \beta \)
Chapter II: Translational profiling of platelets and reticulocytes

Abstract

The mature, circulating forms of the erythroid and megakaryocyte cell lineages are the primary oxygen-carrying and hemostatic cells (or cell fragments) of blood. Young reticulocytes and platelets lack nuclei and do not transcribe new RNA, but inherit a pool of ribosomes and mRNAs that enable protein synthesis during the final phases of maturation. Here, we comprehensively profiled translation by platelets and reticulocytes using ribosome profiling. These results identified more than 6,700 mRNAs translated into protein by two primary cell types in the final phases of terminal differentiation. Surprisingly, this work led us to describe a major novel feature of translation by non-nucleated blood cells that involves the dramatic accumulation of nontranslating ribosomes in the 3’UTRs of platelet and reticulocyte mRNAs.

Introduction

In adult mammals, the primary oxygen carrying and hemostatic blood cell lineages become non-nucleated through the process of terminal differentiation (Fig. 1a). The reasons for this unique feature of mammalian development are still unclear. However, at least for red blood cells, it is widely believed that nuclear elimination permits the deformability required to pass through narrow vessels of the microcirculation. The reasons for non-nucleated platelets in mammals, while all other vertebrates have
nucleated “thrombocytes” that serve the hemostatic function of blood, is unknown\(^2\).

The presence of long-lived mRNAs and ribosomes in platelets was discovered in the 1960s and 70s\(^4,5\), followed shortly after by the first descriptions of platelet protein synthesis\(^6-8\). Recently, high throughput approaches revealed unanticipated complexity to the platelet transcriptome\(^9-12\), and several specific examples of translation and translational control in platelets are reported in the literature\(^13-18\). In erythroid progenitors, chromatin condensation and transcriptional arrest precede enucleation at the late erythroblast stage, during which time the translation of hemoglobin mRNA is highest\(^3,19\). Exceptionally stable RNAs or protein synthesis by non-nucleated or transcriptionally arrested tissues is well documented and occurs in lens\(^20\), oocyte fragments\(^21\), mud snails\(^22\), algae\(^23\), and plants\(^24\).

Here, we comprehensively profiled translation by primary human platelets, reticulocytes, and several in vitro cell culture-based models of platelet or red blood cell formation using ribosome profiling, which relies on the deep sequencing of ribosome protected mRNA fragments (RPFs)\(^25\). First, our results defined the specific mRNAs translated in platelets and reticulocytes, greatly expanding the known list of proteins synthesized by the mammalian non-nucleated blood cell lineages. These data help to reconcile questions about the contribution of platelet protein synthesis to the platelet proteome\(^26-34\). Second, our results revealed a global accumulation of ribosomes in the 3’UTRs of platelet and reticulocyte mRNAs. Further analysis suggests 3’UTR-bound ribosomes are vacant (lacking a nascent polypeptide or tRNAs) 80S complexes that accumulate during the process of differentiation. Finally, using a comparative RNA

---

\(^2\) A long-held view that platelets are required for the placental birthing mechanism that occurs in mammals was recently disproven. See Michelson A. *Platelets* 2013.
sequencing approach of both megakaryocytes and platelets, we discovered that the platelet transcriptome consists of only a subset of all potential megakaryocyte transcripts, and thus raises the question of how specific mRNAs are selected for platelet loading. We propose a transcript selection model that works through either specific sorting/trafficking of megakaryocyte mRNAs to nascent proplatelets, or through selective mRNA degradation/stabilization by platelets, and present evidence supporting a role for the fragile-X mental retardation protein (FMRP) in this process.

Results

Comprehensive profiling of translation by platelets reticulocytes
Ribosome profiling provides a genome-wide view ribosome occupancy through the deep sequencing of ribosome-protected mRNA fragments\textsuperscript{25}. In this approach, polysomes are released from cells and stabilized by addition of cycloheximide, an inhibitor of translation. RNase is then added to the sample to digest unprotected mRNA, leaving only short, ~28 nucleotide fragments of mRNA which are physically protected by the 80S ribosome (“monosomes”). Monosomes, and their associated mRNA “footprint” are pelleted through a sucrose cushion and the resulting RNA is used to generate a deep sequencing library\textsuperscript{35}. Using ribosome profiling, we identified >6,700 mRNAs that are translated into protein by platelets (Fig. S3a). This large group of mRNAs encode proteins with diverse functions, but are enriched in those encoding cytokines and chemokines involved in the early events of innate immune responses (Fig. S5a), consistent with the evolutionary origin of platelets as immune effector cells\textsuperscript{1,36}. High
throughput sequencing of platelet RNA, which carries greater sensitivity to detect low abundance mRNAs, also identified >10,000 protein coding transcripts present in platelets (Fig. S4c), consistent with transcriptome analyses published by other groups\textsuperscript{9-12}. Prior to this work, only 16 proteins had been reported in the literature as being synthesized in platelets (Fig. S4a). Thus, our work represents a significant expansion of the catalogue of mRNAs translated into protein by platelets and provides evidence that platelets synthesize important chemokines/cytokines involved in inflammation and innate immunity, perhaps suggesting what the role of platelet protein synthesis may be \textit{in vivo}.

In contrast, our primary reticulocyte samples were dramatically enriched in RPFs derived from hemoglobin mRNAs, consistent with the reticulocyte stage of erythroid differentiation being the primary site of hemoglobin synthesis\textsuperscript{19}. RPFs derived from mRNAs encoding adult hemoglobin genes (\textit{HBA1, HBA2, HBB,} and \textit{HBD}) comprised 97.1% +/- 0.1% of all RPFs present in each dataset (Fig. S3b). Using a commercially available hemoglobin RNA-depletion strategy that separates and removes RNA fragments which are complementary to the hemoglobin gene from purified total RNA samples, the contribution of globin-derived RPFs could be reduced to 66.7% +/- 2%. Using this strategy, we identified ~400 additional non-hemoglobin mRNAs translated in reticulocytes.

\textbf{Abundant RPFs in 3’UTRs of platelets and reticulocytes}

Surprisingly, our ribosome profiling datasets from platelets and reticulocytes showed increase (>30-fold) in RPFs mapping to 3’UTRs compared to other nucleated cells types such as Meg01 cells or K562 cells (Fig. 2). To quantify this effect, we used a measure
that reports the relative abundance of RPFs in the 3’UTR compared to the coding sequence (CDS) of a given mRNA in order to control for mRNA abundance and overall translation. Relative 3`UTR occupancy is, therefore, the ratio of length-normalized RPF counts in the 3´UTR to the length-normalized RPF count in the CDS for each mRNA.

\[
\text{Relative 3’UTR ribosome occupancy} = \frac{\text{Total RPFs}_{\text{3’UTR}} / \text{Length of 3’UTR (nt)}}{\text{Total RPFs}_{\text{CDS}} / \text{Length of CDS (nt)}}
\]

Using this measure, we determined that mRNAs translated in platelets or reticulocytes were globally affected by an increase in 3´UTR occupancy (Fig. 2a-b), with transcripts at the 50\(^{\text{th}}\) percentile of 3´UTR occupancy having >30-fold more 3´UTR RPFs after normalization for overall expression level. This indicates that, a typical mRNA in platelets or reticulocytes has >30-fold more 3´UTR RPFs than a typical mRNA in another cell type, such as Meg01 cells or K562 cells.

We next asked whether \textit{in vitro} tissue culture based models of platelet and red blood cell development might recapitulate this finding. To address this question, we purified platelet-like particles (PLPs)\(^{37-40}\) from the cell culture media of megakaryoblastic leukemia cells (hereafter Meg01 cells) (Fig. S2) and prepared ribosome profiling libraries from both PLPs and Meg01 cells. Interestingly, PLPs showed a similar, >30-fold enrichment in 3´UTR RPFs relative to Meg01 cells for typical mRNAs (Fig. 3a), consistent with our observations in primary human platelets. We also prepared ribosome profiling libraries from erythroleukemia (K562) cells\(^{41}\) during erythroid differentiation induced towards the erythroid lineage with the protorphorpyrin molecule hemin\(^{42-45}\). Strikingly, 3´UTRs were progressively enriched >30-fold in RPFs during
differentiation of K562 cells to an extent similar to that observed in primary reticulocytes (Fig. 3b-c).

3′UTR RPFs result from unrecycled post-termination ribosomes

To further understand the nature of 3′UTR RPFs we observed in our ribosome profiling data, we analyzed the RPF read length distribution of RPFs mapped to coding sequences and 3′UTRs. RPF size distributions for elongating ribosomes in the CDS and for RPFs mapped to the 3′UTR were nearly indistinguishable in PLPs, indicating 3′UTR RPFs are unlikely to be the result of non-specific ribonucleoprotein (RNP) complexes which might conceivably protect short stretches of mRNA as the ribosome does46 (Fig. 5b).

Ribosomes could potentially enter the 3′UTR as the result of several distinct processes including translational read-through of stop codons47, alternative open reading frame usage48, or a failure to recycle subunits after peptide release49,50. To discriminate among these possibilities, we performed several analyses. First, we analyzed the reading frame distribution for RPFs mapped to coding sequences or 3′UTRs. In coding sequences, 74% and 67% of RPFs occupied a single open reading frame in Meg01 cells and PLPs, respectively (Fig. 5a). However, in 3′UTRs, no reading frame bias was evident and RPFs occupied all three potential reading frames similarly. This suggests ribosomes in PLPs are not the result of translational read-through and are likely not engaged in translation. However, loss of reading frame earlier in the coding sequence could, in principle, lead to 3′UTR translation in variety of reading frames, obscuring a reading frame signal while still representing a signature of elongating (thus translating) ribosomes. Thus, we next tested whether ribosomes in the 3′UTR are associated with a
nascent polypeptide. Importantly, release of nascent polypeptide chains from mammalian polysomes results in 80S ribosomes that are readily dissociated into subunits by high salt washes\textsuperscript{51-55}. Taking advantage of this fact, we prepared ribosome profiling libraries from PLP polysomes treated with a stringent (1M KCl) high salt wash followed by buffer exchange and nuclease footprinting. As validation for this approach, our high salt washed samples were strongly depleted for terminating ribosomes (RPFs containing a stop codon in the A site) (Fig. 4a, black arrows), consistent with normal recognition of the termination codon by eukaryotic release factors and subsequent release of the nascent polypeptide. This indicates the high salt wash is an effective tool to discriminate between elongating ribosomes and other states of the 80S ribosomes such as post-termination ribosomes. Moreover, normal recognition of annotated stop codons during translation rules out the unlikely possibility that a massive change, or failure, in RNA splicing leads to translation into 3’UTRs. Importantly, the high salt wash strongly dissociated 3’UTR ribosomes (Fig. 4b-c), indicating that the majority of 3’UTR ribosomes in PLPs are likely not engaged in translation. Interestingly, the RPF size distribution in high salt washed sample libraries was narrowed, and ribosomes primarily protected a 27nt mRNA fragment (Fig. 5b, right panel). This suggests more stringent conditions may favor a uniform conformation of the ribosome, perhaps through ejection of labile E site tRNAs\textsuperscript{54, 56} or stabilization of the L1 stalk\textsuperscript{57, 58}. Consistent with this interpretation, our high salt wash led to an apparent improvement in reading frame signal in coding sequences (Fig. 4a), which likely reflects stabilization of a single conformational state of the ribosome, as has been observed for drugs that target translation\textsuperscript{59}.
**Translational control during platelet activation**

As platelets are non-nucleated cell fragments, incapable of synthesizing DNA or RNA, we tested whether platelet activation is associated with the control of gene expression at the level of translation as has been reported by others\(^\text{13}\). Using ADP or thrombin – both physiological activators of platelets – we stimulated platelets *ex vivo* and measured surface expression of the activation-dependent marker protein CD62 using flow cytometry (Fig. S9a-c). After 60 minutes, we measured changes in the translation of specific mRNAs by ribosome profiling. We observed modest changes in the translation of specific mRNAs following activation with either ADP or thrombin (Fig. S11a-c).

After analysis of differential footprint count changes, we identified 30 mRNAs that were translationally stimulated (1.4 to 6.4 fold) as well as 36 mRNAs whose translation decreased (1.7 to 5 fold) (Fig. S10a). Among those translationally stimulated were mRNAs encoding the major platelet chemokines platelet factor-4 (PF4 or CXCL4) and CCL5 (*RANTES*) which are both involved in recruiting neutrophils, monocytes, and lymphocytes to sites of inflammation\(^\text{60}\).

**Correlation between the platelet transcriptome, translatome, and proteome**

To assemble a genome-wide view of MK and platelet gene expression, we analyzed relationships between megakaryocyte and platelet mRNA abundance, translation, and protein level using available proteomics datasets\(^\text{27, 30, 61-63}\) (Fig. S6a). We found that platelet RNA abundance correlates with protein abundance only weakly (spearman’s rho=0.2), consistent with the reports of others\(^\text{26, 64, 65}\), and this correlation is significantly improved by analyzing ribosome occupancy instead of mRNA abundance (rho=0.65)
(Fig. S6a-b). This indicates that protein synthesis by platelets helps shape the platelet proteome. Surprisingly, we found that mRNA abundance in platelets – after normalizing for expression level in megakaryocytes – still varied by over three orders of magnitude (Fig. S12), with some mRNAs being strongly enriched in platelets relative to their abundance in megakaryocytes. This suggests the platelet transcriptome undergoes a process of selection either due to specific mRNA trafficking/sorting into nascent platelets, reminiscent of neuronal mRNA localization in dendrites, or selective stabilization/ degradation of mRNA subsets once present in platelets. Finally, analysis of platelet proteomics data revealed that ribosomal proteins are globally depleted in platelet datasets compared to other tissues (Fig. S8, red points), consistent with protein synthesis occurring in only a fraction of “young” platelets in circulation.

**Role for FMRP in MK/platelet mRNA trafficking or translation**

With our list of strongly enriched platelet mRNAs (>32 fold enriched, n=60), we used de novo motif identification software (MEME) to identify a set of 3’UTR sequence motifs that correlated with mRNA enrichment in platelets (Fig. S14). Next, we cross-referenced those motifs with known target sites of RNA binding proteins curated by the RNA Binding Protein Database (RBPDB). From these results, the *FMR1* gene emerged as a promising candidate given its role as a putative translational regulator during mRNP trafficking in neurons. Interestingly, FMR1 mRNA is present in megakaryocytes and FMRP protein is found at the sites of nascent proplatelet formation in Meg01 cells (Fig. S15). The motif we identified (Motif 2 in Fig. S13a) reportedly forms a G-quadraduplex structure capable of binding FMRP with high affinity in vitro.
and several platelet mRNAs, for example SPARC, contain tandem copies of Motif 2, (Fig. S13c). We therefore hypothesized that FMRP might regulate the translation of a subset of mRNAs in MKs which are trafficked into nascent platelets, consistent with the putative role for this protein in neurons/dendrites.

Next, we tested the effect of transient knockdown of FMRP in Meg01 cells by transfecting FMR1-targeted siRNAs or control duplexes. After 5 days, knockdown of FMRP led to significant changes in the translation of hundreds of mRNAs using a strict false-discovery rate of 0.01% (Fig. S16b-d). A modest correlation (rho=−0.2) existed between mRNAs found enriched in PLPs and those down regulated after FMRP knockdown (Fig. S16 continued). Affected mRNAs were not enriched in the motif identified in primary platelets, perhaps reflecting transcriptome differences between primary platelets and Meg01-derived PLPs. This result, however, still suggests that in Meg01 cells, FMRP may regulate the translation of a class of mRNAs that are trafficked into nascent platelets. The mechanism involved, however, is unclear.

Finally, to test whether loss of FMRP in vivo results in a functional defect in bone marrow megakaryocytes, we prepared bone marrow aspirates of FMR1-null mice, or littermate controls, and analyzed them by electron microscopy. Megakaryocytes are identified by their large size (25-100 µm), multiple nuclei, and demarcation membrane system (DMS). FMR1-null bone marrow megakaryocytes had fewer demarcation membranes and markedly asymmetric alpha granule distributions (Fig. S17), indicating less mature, non-platelet producing, megakaryocytes. No overt hematologic phenotype has been reported in fragile-X syndrome patients, however platelet defects
have been observed\textsuperscript{80}. Compensation for loss of \textit{FMR1} could occur through up
regulation of \textit{FMR1} paralogues (\textit{FXR1}, \textit{FXR2})\textsuperscript{81}.

\textbf{Discussion}

Prior to this work, only a limited number (sixteen that were readily identified in the
literature) of specific examples of proteins synthesized by platelets were known (\textbf{Fig. S4b}). The results presented in this chapter greatly expanded the number of mRNAs
known to be translated in platelets (by more than 6,700 mRNAs). Most significantly,
however, we have provided the first evidence that translation termination and recycling
may be specifically regulated by blood cell lineages (Further detailed in Chapter III).

\textit{Evidence for nontranslating 80S ribosomes in the 3’UTR}

Our ribosome profiling experiments indicate that post-termination ribosomes accumulate
in 3’UTRs in platelets and reticulocytes. Several lines of evidence support the view that
these are \textit{bona fide} ribosomes which first translate the coding sequence, release the
polypeptide, then enter the 3’UTR after failing to dissociate from the mRNA. First, the
80S ribosome protects a unique \textasciitilde28-32 nucleotide mRNA fragment\textsuperscript{84-86}, which in
ribosome profiling libraries leads to a highly reproducible RPF size distribution\textsuperscript{25, 35, 46, 87}
. In our data, the RPF size distribution in 3’UTRs followed exactly this distribution and
was nearly indistinguishable from RPFs mapped to coding sequences. Second, a high salt
wash of polysomes led to a reduction in 3’UTR RPFs, consistent with the behavior of
post-termination ribosomes, lacking a nascent polypeptide chain. Third, our high salt
wash experiments led to a overall shortening of ribosome footprints by 1 nucleotide on
the 3’ end of each sequencing read, possibly due to ejection of the E-site tRNA\textsuperscript{54,56}. While most 3´UTR footprints were cleared by a high salt wash, those remaining also underwent a shortening by 1 nucleotide, suggesting they react similarly to high ionic strength as elongating ribosomes in coding sequences (Fig. 5). Additional specificity for 80S ribosomes could be obtained through the use of tagged 60S ribosomal proteins\textsuperscript{46} or specific release by over expression of ribosome recycling factors (See Chapter III).

Additional evidence makes it unlikely that ribosomes initiate \textit{de novo} in 3´UTRs. 3´UTR ribosome abundance strongly correlates with the number of ribosomes occupying the coding sequence. Moreover, in eukaryotes, translation initiation is a highly regulated process that requires the 5´ end of the mRNA molecule. Widespread initiation of ribosomes in the 3´UTR, therefore, is exceedingly unlikely.

\textbf{Are platelet/reticulocyte mRNAs in the process of being degraded?}

Messenger RNAs are typically unstable (relative to protein), presumably to allow for the regulation of gene expression. The rate determining step of mRNA decay is often deadenylation, and is, thus, often followed by rapid clearance of the message\textsuperscript{89,90}. This process usually does not give rise to detectable decay intermediates in wild type backgrounds (unless additional genetic changes are made which block downstream mRNA decay, for example deletion of XRN1 or exosome components). In these artificial genetic backgrounds, ribosomes have been described which translate to the 3´ end of an endonucleolytically cleaved mRNA protect a characteristically shorter stretch of mRNA ~16-18 nucleotides in ribosome profiling studies\textsuperscript{50}. In our platelet and reticulocyte profiling data, these shorter RPFs comprise only ~1-3% of total RPFs, and no greater a
fraction that in other cell types, indicating a very small fraction of ribosomes are involved in the translation of truncated mRNAs and these fragments are usually not detectable in wild type cells. Consistent with this view, high throughput sequencing has confirmed that platelet mRNA abundance measurements using mRNA isolated by polyA selection are comparable to measurements made by after preparing total RNA (with or without rRNA depletion prior to sequencing). Platelet mRNAs are also known to contain the modified m7G-cap which is found at the 5′ end of eukaryotic mRNAs. Additional experiments in rabbit reticulocyte lysates programmed with full length in vitro transcribed, capped, and polyA-tailed firefly luciferase (FLUC) RNA demonstrated that 3′UTR ribosomes accumulate early after translation begins does not require any intrinsic feature (e.g. secondary structure of sequence motif) of an mRNA to occur. General mechanisms for promoting long-lived mRNAs are not known, but could involve cis or trans acting factors that help mRNAs avoid recognition by mRNA decay pathways. Alternatively, differentiating cells could, in principle, down regulate mRNA decay pathways themselves (See Chapter III).

Materials and Methods

Primary blood products Human platelets were obtained from plateletpheresis blood samples from healthy volunteers. Samples were collected by Akron Biotech or Research Blood Components using protocols approved by the institutional review board (IRB) and described in IRB #00001463 (Johns Hopkins University). Primary peripheral blood
mononuclear cells and primary bone marrow mononuclear cells were obtained from ATCC.

**Platelet purity** To confirm the purity of our platelet samples, which are often contaminated by mononuclear cells such as white blood cells, we analyzed our plateletpheresis samples by flow cytometry, which because of the small size of platelet particles (0.1-2µm in diameter), can easily detect the presence of contaminating mononuclear cells such as white blood cells (Fig. S1a). After collecting 50,000 events on the flow cytometer, we did not observe any particles of a size compatible with being a nucleated cell. We also fixed and immuno-stained plateletpheresis samples for expression of CD42p, a constitutive platelet marker that encodes the surface-exposed heterodimer glycoprotein Ib (GP1BA/B), and found 97.1% of platelet-sized particles were CD42p-positive compared to an IgG control (Fig. S1b). Finally, our ribosome profiling datasets were strongly depleted of the common leukocyte antigen (CLA; CD45, aka PTPRC) (Fig. S3a, green point). Taken together, this suggests our plateletpheresis samples do not contain significant amounts of contaminating mononuclear cells.

**Tissue culture of human cell lines** Meg01 and K562 purchased from ATCC and cultured in RPMI 1640 (10% FBS). K562 media was supplemented with 2mM glutamine. All experiments were performed with cells during passage numbers 5-25.

**Platelet-like particle (PLP) culture** To obtain PLPs, semi-adherent Meg01 cells were grown to 80% adherent confluence and culture supernatants were aspirated. Suspension
cells were pelleted at 100xg for 5 minutes at room temperature. Post-cell supernatants were aspirated and discarded. Cell pellets were resuspended in 1ml cell culture media and added back to adherent Meg01 cells with fresh media. Meg01 cells were returned to the 37°C tissue culture incubator and allowed to release PLPs for indicated amount of time (ranging from 24 hours to 7 days). 1:10,000 dilutions of activating anti-FasL (Millipore) was added to Meg01 cell culture to increase PLP production.

**Ribosome profiling library generation** RPF libraries were generated as previously described\(^\text{35}\). For primary human platelets and PLPs, 0.1-1µg of total RNA was used as starting material; for human cell line samples, 10-20µg total RNA was used.

**PLP high salt wash experiments** For high salt washes, clarified PLP sample lysates were mixed with an equal volume of 2M KCl in polysome lysis buffer (PLB) resulting in a final KCl concentration of 1M. In control samples, an equivalent volume of polysome lysis buffer, but without supplemented KCl, was added to sample lysates, ensuring identical lysate concentrations during the subsequent incubation. Lysates were incubated on ice for 30 minutes. After the incubation, the lysate buffer was exchanged using an S400 Zeba desalting column (Pierce). RNaseI digests were then performed as described previously\(^\text{35}\).

**In-lysate RNA quantification** RNA amount was quantitated from clarified lysates using the Quant-it kit (Life Technologies). Samples and provided standards were measured in duplicate using the FluorStar microplate fluorometer (OptiStar).
**RNA-Seq** RNA-Sequencing libraries were prepared using the low-throughput TruSeq stranded total RNA sample prep kit (Illumina) with 0.1-0.5µg total RNA as starting material.

**Platelet and PLP flow cytometry** Platelets and PLPs were washed in sterile 1X PBS and fixed for 15 minutes in 4% para-formaldehyde (PFA) in the dark. Primary antibodies used were FITC-anti-CD62 (Life Technologies), PE-anti-CD41 (Life Technologies), and FITC-anti-CD42 (Life Technologies).

**Meg01 cells immunofluorescence** Meg01 cells were grown on poly-lysine coated cell staining chambers until ~70% confluence. Adherent cells were washed and fixed in 4% PFA for 15 minutes in the dark. Primary antibodies used were anti-FMRP (abcam) and anti-Tubulin (abcam), and phalloidin (Life Technologies).

**FMR1-null mice and electron microscopy** FMR1 knockout mice were obtained from Jackson Laboratories. All mice used were between 6 and 8 weeks old and male. For EM studies, mice were sacrificed by cervical dislocation and femurs were dissected and fixed overnight at 4°C. Electron microscopy was performed at the Carnegie Institution of Washington Department of Embryology (Baltimore, MD).

**High-throughput sequencing** All sequencing was performed using the core facilities at either the Carnegie Institution of Washington Department of Embryology (Baltimore,
MD), McKusick-Nathans Institute for Genetic Medicine at the Johns Hopkins Medical Institutions (JHMI), or the University of California Berkeley. All sequencing was performed using either a HiSeq2000 or HiSeq2500 instrument (Illumina).

**Figure List**

Figure 1. Hematopoietic cell lineage map

Figure 2. Platelets and reticulocytes accumulate 3´UTR ribosomes

Figure 3. 3´UTR ribosomes accumulate during *in vitro* differentiation of Meg01 and K562 cells

Figure 4. 3´UTR ribosomes are not translating a polypeptide

Figure 5. Ribosomes in the 3´UTR are non-translating post-termination ribosomes

Figure S1. CD42p expression in plateletpheresis samples

Figure S2. CD41 expression by Meg01-derived platelet-like particles (PLPs)

Figure S3. RPF scatter plots of primary platelets and CD71-selected reticulocytes

Figure S4. Proteins previously reported to be synthesized in platelets

Figure S5. Gene ontology of highly expressed platelet mRNAs

Figure S6. Correlation among global gene expression measurements in platelets

Figure S7. Translational efficiency (TE) of platelet mRNAs

Figure S8. Ribosomes are depleted in platelets

Figure S9. Platelet activation by thrombin or ADP

Figure S10. Translational changes following platelet activation

Figure S11. RPF scatter plots of activated and resting platelets

Figure S12. Many platelet mRNAs are enriched relative to megakaryocytes
Figure S13. *De novo* motif searches of platelet-enriched mRNAs

Figure S14. Other identified RNA motifs and interacting-RBPs

Figure S15. FMRP is expressed by Meg01 cells

Figure S16. FMRP is important for expression of a subset of PLP-enriched mRNAs

Figure S17. *FMR1*-null mice have abnormal megakaryocyte differentiation

**Figure Legends**

**Figure 1. Hematopoietic cell lineage map** | **a.** Blood cell lineages that give rise to a non-nucleated terminally differentiated form (light blue oval) are indicated. HSC-derived mononuclear cell types are also indicated (pink oval).

**Figure 2. Platelets and reticulocytes accumulate 3´UTR ribosomes** | **a.** Empirical cumulative distribution function (ECDF) plots of relative 3´UTR ribosome occupancy in primary human platelets (dark and light red lines; replicates of a single donor’s sample) or Meg01 cells (black line). 3´UTR occupancy was calculated as the ratio of 3´UTR RPF density / CDS RPF density for each mRNA passing expression threshold (See methods).  
**b.** ECDF plots of relative 3´UTR ribosome occupancy for primary human CD71+ reticulocytes (dark and light red lines; replicates) or K562 cells (black).  
**c.** Representative ribosome profiles for specific transcripts indicated. Thick blue boxes designate coding sequences and thin blue boxes designate 5´- or 3´UTRs. Yellow boxes highlight 3´UTRs.
Figure 3. 3’UTR ribosomes accumulate during in vitro differentiation of Meg01 and K562 cells | a. ECDF plots of relative 3’UTR ribosome occupancy for Meg01-derived platelet-like particles (PLPs) (dark blue) or precursor Meg01 cells (black). b. ECDF plots of relative 3’UTR ribosome occupancy for naïve K562 cells (black) or differentiated K562 cells induced for 2 days (dark blue) or 10 days (light blue) with 50 µM Hemin. c. mRNA-specific ribosome profiles show 3’UTR ribosome accumulation only in differentiated K562 cells.

Figure 4. 3’UTR ribosomes are not translating a polypeptide | a. Metagene plots of RPFs near termination codons from control PLPs (dark blue) or after a high salt wash (1M KCl) (pink) (See methods). Arrows indicate where terminating ribosomes can be identified with the stop codon in the A site of the ribosome. The high salt wash releases ribosomes not stabilized by the presence of a nascent polypeptide in the ribosomal exit tunnel. b. ECDF of 3’UTR ribosome release following the high salt wash. “Sensitive” transcripts reflects those mRNAs whose 3’UTR ribosomes were most effectively released by the high salt wash using an arbitrary threshold of 4-fold clearing (gold box). c. mRNA-specific ribosome profiles demonstrating high salt release of 3’UTR ribosomes in PLPs (Meg01 cells: black; PLP-control: dark blue; PLP-high salt: pink).

Figure 5. Ribosomes in the 3’UTR are non-translating post-termination ribosomes | a. Unlike elongating ribosomes, found in coding sequences, 3’UTR-bound ribosomes do not exhibit a collective, detectable reading frame bias (Meg01: black; PLP-control:dark blue; PLP-high salt: pink) Reading frames are designated f0 (for the reading
frame defined by the protein’s coding sequence), f1 (+1 frameshifted), and f2 (+2 frameshifted). b. RPF size distributions plotted versus the fraction of total read number (Meg01: black; PLP-control: dark blue; PLP-high salt: pink) for coding sequences (solid lines) and 3’UTRs (dotted lines).

**Figure S1. CD42p expression in plateletpheresis samples** | a. Plateletpheresis samples were analyzed for forward (FSC.A) and side (SSC.A) scatter properties by flow cytometry with comparison to standard beads (1µm and 3 µm in diameter). b. Platelets were fixed and stained with a FITC-conjugated anti-CD42p primary antibody (a constitutive platelet marker protein).

**Figure S2. CD41 expression by Meg01-derived platelet-like particles (PLPs)** | a. Histogram analysis of forward scatter (cell size) of Meg01 cells and supernatants containing platelet-like particles (PLPs) obtained by flow cytometry. b. PLPs were collected over 24 hours and purified by centrifugation and filtration through a 5µm filter. c. PLPs were fixed in 4% PFA and stained with a PE-conjugated anti-CD41 primary antibody for flow cytometry. d. Histrogram depicting PE-anti-CD41 staining of PLPs.

**Figure S3. RPF scatter plots of primary platelets and CD71-selected reticulocytes** | a. Ribosome profiling replicates for primary human platelet samples. Platelet marker genes are indicated in blue. PTPRC (CD45; common leukocyte antigen) is indicated in green. b. Ribosome profiling replicates for primary CD71+ human reticulocytes. Hemoglobin mRNAs are indicated in red.
**Figure S4. Proteins previously reported to be synthesized in platelets** | a. Proteins previously reported as being synthesized by human or mouse platelets with literature references provided. For comparison, platelet RNA abundance measurements reported in the present study as well as those from RNA-Seq data recently published by Rowley *et al.* is provided. Ribosome profiling reads (current study), proteomic data (published sources), and primary human megakaryocyte RNA-Seq data (Nürnberg *et al.*) is also indicated. b. Venn diagram indicating the expansion of the platelet translat-ome. c. Total numbers of mRNAs, proteins, and translated RNAs in human platelets.

**Figure S5. Gene ontology analysis of highly expressed platelet mRNAs** | a. Platelet mRNAs with the highest RPF densities were rank ordered and provided to the Gorilla gene ontology analysis program ([http://cbl-gorilla.cs.technion.ac.il/](http://cbl-gorilla.cs.technion.ac.il/)) using a single ranked list of genes.

**Figure S6. Correlation among global gene expression measurements in platelets** | a. Spearman correlations between global gene expression measurement methods in megakaryocytes and platelets. Values reported are spearman’s rho. b. Platelet genes were binned using protein abundance (ppm) measurements and used to generate box and whisker plots of RPF density.
Figure S7. Translational efficiency (TE) of platelet mRNAs | Platelet mRNA translational efficiencies (TE) were plotted against RNA abundance (RPKM) (a) or protein abundance (ppm) (b).

Figure S8. Ribosomes are depleted in platelets | a. Scatter plots of platelet protein abundance (ppm) versus protein abundance in other human tissue types. Ribosomal proteins are indicated in red.

Figure S9. Platelet activation by thrombin or ADP | a. Schematic illustrating how platelets can be activated ex vivo by the addition of the serine protease thrombin. The oligopeptide Gly-Pro-Gly-Arg (GPPG) is used as a competitive inhibitor of fibrin polymerization b. Platelets were activated with either thrombin (0.1U/ul) or ADP (0.5µM) for 30 minutes and then fixed and stained using a FITC-conjugated anti CD62p primary antibody (a marker for platelet activation). c. Flow cytometry data demonstrating both an increase in CD62p surface expression following thrombin addition as well as dramatic change in morphology reflecting platelet degranulation.

Figure S10. Translational changes following platelet activation | a. Resting or activated platelets were used for ribosome profiling experiments. Venn diagram indicating the number of mRNAs whose translation was significantly increased (pink) or decreased (blue) in activated versus resting samples. Inset shows gene ontology analysis of significant genes that underwent translational stimulation.
Figure S11. RPF scatter plots of activated and resting platelets | Replicate scatter plots of RPFs mapping to coding sequences in resting (a), ADP-stimulated (b), or thrombin activated (c) human platelets.

Figure S12. Many platelet mRNAs are enriched relative to megakaryocytes | a. Using megakaryocyte and platelet RNA-Seq data, each mRNAs relative platelet enrichment was calculated as $\frac{\text{RPKM}_{\text{platelets}}}{\text{RPKM}_{\text{Megakaryocytes}}}$. b. Replicate enrichment scores calculated from Meg01 and PLP mRNA-Seq data.

Figure S13. RNA sequence motifs found in platelet-enriched mRNAs | a. Using the top 60 most enriched mRNAs in platelets relative to megakaryocytes, several short RNA motifs were identified. b. These motifs are predicted to have the potential to participate in G quardaduplex structures between single or multiple RNA molecules. c. One highly enriched mRNA in platelets relative to megakaryocytes in SPARC, which also contains tandem Motif2 sites in its 3’UTR.

Figure S14. Other identified RNA motifs and interacting-RBPs | a. Other RNA sequence motifs identified from platelet-enriched mRNA sequences as well as RNA binding proteins that have been reported to interact with sequences closely resembling those identified.

Figure S15. FMRP localizes to the nascent platelet region of Meg01 cells | a. Immunofluorescence staining of Meg01 cells for FMRP.
Figure S16. FMRP is important for expression of a subset of PLP-enriched mRNAs | Meg01 cells were transiently transfected with siRNA duplexes targeting FMR1 mRNA and harvested for ribosome profiling after 5 days along with cells treated in parallel with a control siRNA duplex. a. CDS RPF replicate scatter plot of Meg01 cell samples transfected with control siRNA duplexes. b. CDS RPF scatter plot of the FMR1 knockdown sample versus control. c. Differential expression analyses were performed using DESeq R package. Fold changes in RPFs in knockdown versus control cells were plotted against the sum of RPFs in both conditions (MA plot) with significantly differentially expressed mRNAs indicated in blue (FDR=0.01%). d. Volcano plot illustrating adjusted p-values for differentially expressed mRNAs (blue) and all mRNAs (gray). e. Fold changes in RPFs after FMR1 knockdown versus PLP RNA enrichment score relative to Meg01 cells. Differentially expressed mRNAs are indicated in blue.

Figure S17. FMR1-null mice have abnormal megakaryocyte differentiation | a. Bone marrow aspirates from male FMR1-null mice were analyzed by electron microscopy (EM). Megakaryocytes (MKs) were visually identified as large, polyploids cells with demarcation membranes (DMS) indicating nascent platelet “territories”. FMR1-null MKs had immature DMS and asymmetrically distributed alpha granules.
References


60. Weber, C. Platelets and Chemokines in Atherosclerosis.


Fig. 2

(a) Meg01 Platelets (replicates)

(b) K562 Reticulocytes (replicates)

(c) Platelets

- HMGB1

- CALM2

Reticulocytes

- S100A4

- IFI27

Nucleotides of mRNA
Fig. 3

(a) ECDF of 3' UTR occupancy for Meg01 PLPs, HNRNPA1 (-0.6) and HSP90AA1 (-1.23).

(b) ECDF of 3' UTR occupancy for K562, K562+Hemin (2d), K562+Hemin (10d), EEF2 (-1.2), LDHA (-3.66), EEF2 (-7.41), EEF2 (-6.3), and LDHA (-7.88).

(C) RPFs for LDHA and EEF2 under Uninduced, Hemin (2d), and Hemin (10d) conditions.
Fig. 4

(a) Histograms showing 39mer and 27mer RPFs before and after treatment with PLPs and 1M KCl wash. Termination ribosome and post-termination ribosome release markers are indicated.

(b) Graph showing the fraction of transcripts versus fold 3’UTR ribosome clearance.

(c) Heatmaps of RPFs for HSP90AA1 and HNRNPA1 under different conditions: Meg01, PLP-control, and PLP-high salt. Nucleotides of mRNA are indicated.
Fig. S2

(a) Plot P04, gated on P02.R5.R1

(b) 5μm filtered PLPs

(c) Plot P04, gated on P01.PLP-SSC

(d) Plot P06, gated on P01.PLP-SSC
Fig. S3

(a) Platelets

(b) Reticulocytes
### Fig. S4

#### a

<table>
<thead>
<tr>
<th>Gene</th>
<th>Reference</th>
<th>Note</th>
<th>RNA-Seq (rpkmm)</th>
<th>Ribo Prof</th>
<th>Protein (ppm)</th>
<th>RNA-Seq (rpkmm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACAT1</td>
<td>Actn</td>
<td>Warscho et al. 2004</td>
<td>2654.5</td>
<td>1246.1</td>
<td>6019.1</td>
<td>21897.1</td>
</tr>
<tr>
<td>ALB</td>
<td>Albumin</td>
<td>H€ffer et al. 1987</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3021.7</td>
</tr>
<tr>
<td>BCL3</td>
<td>Bcl-3</td>
<td>Warscho et al. 1998</td>
<td>1.0133</td>
<td>0.0124</td>
<td>0</td>
<td>0.35422</td>
</tr>
<tr>
<td>GP1BA</td>
<td>GPIb</td>
<td>H€ffer et al. 1987</td>
<td>127.46</td>
<td>15.437</td>
<td>63.140</td>
<td>1607.54</td>
</tr>
<tr>
<td>F73A1</td>
<td>F13A1</td>
<td>H€ffer et al. 1987</td>
<td>1665.1</td>
<td>827.42</td>
<td>308.97</td>
<td>3034.69</td>
</tr>
<tr>
<td>F13G1</td>
<td>COX-1</td>
<td>Evangelista et al. 1989</td>
<td>731.98</td>
<td>187.04</td>
<td>269.01</td>
<td>193.186</td>
</tr>
<tr>
<td>GPIX</td>
<td>GPX1</td>
<td>Lemire et al. 1997</td>
<td>1748.2</td>
<td>217.74</td>
<td>261.84</td>
<td>712.506</td>
</tr>
<tr>
<td>TGFα2</td>
<td>GP1b</td>
<td>H€ffer et al. 1987</td>
<td>550.60</td>
<td>46.954</td>
<td>41.463</td>
<td>440.70</td>
</tr>
<tr>
<td>TGFB3</td>
<td>TPαRb</td>
<td>H€ffer et al. 1987</td>
<td>753.96</td>
<td>304.73</td>
<td>44.947</td>
<td>2439</td>
</tr>
<tr>
<td>IL-18</td>
<td>IL-18</td>
<td>Lindemann et al. 2001</td>
<td>0.9102</td>
<td>0.0948</td>
<td>0</td>
<td>2.91318</td>
</tr>
<tr>
<td>SERPIN5</td>
<td>PAI-1</td>
<td>Degrè et al. 2004</td>
<td>11739</td>
<td>1.4154</td>
<td>1.9208</td>
<td>56.2510</td>
</tr>
<tr>
<td>ARH504A</td>
<td>RhoGDI</td>
<td>Warscho et al. 2004</td>
<td>18.500</td>
<td>2.7900</td>
<td>1.2118</td>
<td>678.0</td>
</tr>
<tr>
<td>ARH504B</td>
<td>RhoGDIb</td>
<td>Warscho et al. 2004</td>
<td>222.68</td>
<td>460.92</td>
<td>393.54</td>
<td>657.0</td>
</tr>
<tr>
<td>F3</td>
<td>Thrombin Factor</td>
<td>Prasse et al. 2007</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VWF</td>
<td>VWF</td>
<td>H€ffer et al. 1987</td>
<td>209.05</td>
<td>90.97</td>
<td>38.892</td>
<td>3486.0</td>
</tr>
<tr>
<td>16</td>
<td>VWF</td>
<td>H€ffer et al. 1987</td>
<td>15.608</td>
<td>3.501</td>
<td>0.2942</td>
<td>354.0</td>
</tr>
</tbody>
</table>

#### b

**Translated Platelet RNAs**

- Previously reported: 16
- 6,712 RNAs

#### c

**Proteomics**

- mRNA-Seq (~4,000 proteins)
- Ribo Prof (10,322 genes)
- Ribo Prof (6,712 translated RNAs)

†Burkhardt et al.
Vaudel et al.
Martens et al.
Wang et al.
Fig. S5

a

Cellular Process
- Leukocyte chemotaxis
- Cell chemotaxis
- Pos. reg. of macrophage chemotaxis
- Pos. regulation of platelet activation
- Regulation of granulocyte chemotaxis
- Platelet degranulation

Genes
- PF4V1, PF4, CCL5, PPBP
- PF4V1, PF4, CCL5, PPBP
- PTPRJ, THBS1, CCL5
- SELP, PLEK, PTPRJ
- MPP1, JAM3, PTPRKJ, THBS1, CCL5

Enrichment

Significance
- q-value
  - 0.015
  - 0.001

PFN1, PROS1, SRGN, ITGA2B, PPIA, FLNA, POTEKP, SPARC, PPBP, TIMP1, TUBA4A, SELP, EGF, CD9, ITGB3, CLU, MMRN1, TLN1, LYN, WDR1, PLEK, CD36, VCL, TGFBI, TMSB4X, PF4, CALM3, CAP1, THBS1, APP, ABCC4, CFL1
Fig. S6

a

mRNA-Seq  mRNA-Seq  Ribo Prof  Protein

b

RPF density [log2]

Protein Abundance (ppm)
Fig. S7

(a) Scatterplot showing the relationship between RPKM (log2) and TE (log2).

(b) Scatterplot showing the relationship between Protein abundance (ppm) (log2) and TE (log2).
Fig. S8

a

Protein abundance (ppm) [log2]

ribosomal proteins

Platelets vs. B-Cells

Platelets vs. CD8 T-Cells

Platelets vs. NK Cells

Platelets vs. Liver

Platelets vs. Heart

Platelets vs. Testis

1 10 100 1000

10000

1 10 100 1000

10000

1 10 100 1000

10000

1 10 100 1000

10000

1 10 100 1000

10000
Fig. S10

![Venn diagram showing gene expression categories with fold changes and p-values.]

**Table:**

<table>
<thead>
<tr>
<th>Category</th>
<th>p-value</th>
<th>Enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platelet alpha granule lumen</td>
<td>3.39E-05</td>
<td>45.39</td>
</tr>
<tr>
<td>Vesicle lumen</td>
<td>4.43E-06</td>
<td>41.6</td>
</tr>
<tr>
<td>Cytoplasmic vesicle</td>
<td>2.26E-04</td>
<td>8.4</td>
</tr>
<tr>
<td>Platelet degranulation</td>
<td>3.85E-04</td>
<td>20.38</td>
</tr>
<tr>
<td>Response to cadmium ion</td>
<td>6.52E-04</td>
<td>51.21</td>
</tr>
</tbody>
</table>
Platelet enrichment = \( \frac{\text{RPKM}_{\text{Platelets}}}{\text{RPKM}_{\text{Megakaryocytes}}} \)
Fig. S13

(a) Motif 1 E-value = 3.3 x 10^{-49}

(b) G Quartet

1 strand

2 strands

4 strands

(c) Platelets

Megakaryocytes

5'-UTR

SPARC (uo03lug.3)

3'-UTR

TCTTCTCAGGGCTCAGGGGACTGCCAGGCTGTTTCAGCCAGGAGGGCCAAAATCAAAGA

Motif 2

Motif 2
<table>
<thead>
<tr>
<th>E-Value</th>
<th>Regular Expression</th>
<th>RNA Binding Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3e-049</td>
<td>[CG]AGGC[T]GA[AG]TG</td>
<td>PTBP2, HRNPF, HRNPH1, SFRS1, TRA2A, FMR1, PSi</td>
</tr>
<tr>
<td>1.1e-032</td>
<td>G[CA][CT]ACG[CG][CA][AG][GC]</td>
<td>HRNPF, HRNPH1, HRNPA1</td>
</tr>
<tr>
<td>1.2e-024</td>
<td>G[CA][AG][GG][TG][TG]GG</td>
<td>SFRS1</td>
</tr>
<tr>
<td>3.1e-023</td>
<td>C[TA][GC][AG][CT]AG [AG]AG</td>
<td>PTBP1</td>
</tr>
<tr>
<td>2.2e-023</td>
<td>C[CA][CA][CAT][CA][TA][CA]</td>
<td>PTM1, SFRS1</td>
</tr>
<tr>
<td>3.5e-019</td>
<td>G[GG][CA][AG][CG][AG][GC]C</td>
<td>SFRS1</td>
</tr>
<tr>
<td>2.6e-020</td>
<td>AT[CT][CT][CA][AGA]ATT</td>
<td>PTBP1</td>
</tr>
<tr>
<td>1.0e-016</td>
<td>G[CT][CG][AC][CT][AC][AC]</td>
<td>ELAVL1 [PsUR]</td>
</tr>
<tr>
<td>4.4e-017</td>
<td>GT[GG][CT][CT][CT][CT]GG</td>
<td>ZFP36, ELAVL1 [PsUR], HRNPD, TAL1, KHSRP, JUN1, puf2, puf3, mpo3, AC01, IRES2</td>
</tr>
<tr>
<td>2.5e-018</td>
<td>G[AG][CG][AT][CA][AG]GA</td>
<td>KHSRP</td>
</tr>
<tr>
<td>2.3e-014</td>
<td>G[AG][GC][AC]CAC</td>
<td>SFRS1</td>
</tr>
<tr>
<td>5.4e-011</td>
<td>AGAG[CA][CA][AG][AG]A</td>
<td>y552-0</td>
</tr>
<tr>
<td>3.2e-014</td>
<td>G[TC][CT][CT][CT][CT][CT]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>1.3e-008</td>
<td>C[CA][TA][TA][GG][CG]</td>
<td>PTBP1</td>
</tr>
<tr>
<td>2.7e-012</td>
<td>A[AG][CG][AAA][AAG]A</td>
<td>PTBP1</td>
</tr>
<tr>
<td>5.2e-008</td>
<td>G[CG][CT][CT][CT][CT][CT]</td>
<td>PTBP1</td>
</tr>
<tr>
<td>9.4e-007</td>
<td>C[CGCA][CGA][AT][CC]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>7.8e-005</td>
<td>G[AG][CT][TA][CA][AA][CT]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>4.6e-001</td>
<td>A[AG][TA][CA][AT][TA][A]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>1.1e-005</td>
<td>C[CT][CT][CA][AA][AT][TC]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>7.5e-006</td>
<td>A[GA][AA][AA][ATC][A]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>7.5e-003</td>
<td>CGATCCCAACGC</td>
<td>SFRS1</td>
</tr>
<tr>
<td>9.1e-001</td>
<td>TCCTGAGGCTG</td>
<td>SFRS1</td>
</tr>
<tr>
<td>3.3e-003</td>
<td>C[TA][AT][AA][AGC]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>3.7e-004</td>
<td>GCTTGAAG[AC][GA]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>9.8e-003</td>
<td>CCT[AC][AT][AC][CCA]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>9.5e-002</td>
<td>CTAT[GG][AT][TC]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>5.8e-005</td>
<td>C[CT][GA][CT][AC][TG][GA][AT]</td>
<td>SFRS1</td>
</tr>
<tr>
<td>4.5e-003</td>
<td>A[CT][GA][AA][ATC][A]</td>
<td>SFRS1</td>
</tr>
</tbody>
</table>
Fig. S15

a

FMRP
Tubulin
DAPI

20X

10X

IgG control

▶ proplatelet process
Fig. S16 continued
Chapter III: Regulation of ribosome recycling factors in platelets and reticulocytes

Abstract
Using ribosome profiling, we uncovered a dramatic accumulation of ribosomes in the 3’UTRs of platelet and reticulocyte mRNAs. Here, we describe the underlying mechanism that produces this phenotype. Using in vitro models of erythroid and megakaryoid differentiation, we define a differentiation-dependent switch in the ribosome recycling machinery during early differentiation. This switch involves the loss of canonical ABCE1-mediated ribosome recycling and involves the upregulation of interacting eukaryotic release factor homologues, PELO and HBS1L, that have been previously implicated in mRNA surveillance. We show that PELO and HBS1L are critical for the activation of fetal hemoglobin synthesis during early differentiation of K562 cells and are lost during the late phases of differentiation, leading to abundant 3’UTR ribosomes. Restoring PELO/HBS1L in late differentiated cells releases these ribosomes. Taken together, this work defines physiological roles for PELO/HBS1L beyond mRNA surveillance and revealed an unexpected link to the regulation of fetal hemoglobin expression.

Introduction
The pelota gene was originally identified in 1993 using a P element insertional mutagenesis strategy to identify male sterility mutants in D. melanogaster. Spermatocytes failed to undergo meiotic nuclear reduction, and instead arrested prior to...
the first meiotic division of spermatogenesis\(^1\). The *pelota* locus encodes a 43 kDa protein that is broadly conserved in plants\(^2\), worms\(^3\), yeast\(^1\), archebacteria\(^4\), candida\(^5\), mice\(^6\), and humans\(^7\) and is necessary for meiotic spindle formation as well as breakdown of the nuclear envelope in Drosophila\(^1\). In mice, loss of the *pelota* homologue leads to embryonic lethality\(^8\). In yeast, the *pelota* homologue (*dom34*)\(^3\) mutants were also observed to have meiotic defects\(^9\).

Analysis of the yeast DOM34 sequence led Davis & Engebrecht to describe the homology between DOM34 and the eukaryotic release factor 1 (eRF1) family of proteins involved in translation termination, or the release of the completed polypeptide at the end of translation\(^9,10\). In a screen for high copy suppressors of the *dom34* cell cycle defect, the ribosomal small subunit protein 30A (RPS30A) was identified and provided the first evidence that DOM34 is involved in translation\(^9\). In 2006, Doma and Parker demonstrated that Dom34 and the elongation factor 1 (EF-1\(\alpha\)) paralogue Hbs1 are required in yeast to recognize and degrade aberrant mRNAs that impede the translocation of the ribosome, targeting defective mRNAs for endonucleolytic cleavage and degradation by no-go decay (NGD)\(^11\). Later work showed that Dom34/Hbs1 work as NGD factors, in part, by dissociating (“rescuing”) stalled ribosomes into 40S subunits and 60S subunit-peptidyl-tRNA complexes\(^12-15\), yet the identity of the endonuclease recruited by Dom34/Hbs1 is unknown\(^16\).

Interestingly, Dom34/Hbs1 can also rescue vacant 80S ribosomes (lacking peptidyl-tRNA) *in vitro*\(^13\) and in yeast\(^17\) indicating a role outside of mRNA surveillance. Ribosome profiling studies in yeast revealed Dom34/Hbs1 can rescues post-termination, vacant ribosomes in the 3’UTR\(^18\) after they have failed to be dissociated (“recycled”).

\(^3\) Duplication Of Multilocus region
from the mRNA by ABCE1 (Rli1 in yeast) after releasing the completed polypeptide \textsuperscript{19-21}. Consistent with a role in ribosome rescue, Dom34 restores the growth of cells deficient in ribosomal subunits\textsuperscript{22}. This suggests, albeit indirectly, that stalled ribosomes released by Dom34/Hbs1 may be returned to the cytoplasmic pool for additional rounds of translation.

Here, we demonstrate that the Dom34/Hbs1 homologues in humans, PELO/HBS1L, rescue stalled ribosomes in human cells and are specifically silenced during platelet/red blood cell maturation. This permits the global, tissue-specific accumulation of vacant ribosomes (lacking a nascent polypeptide chain or tRNAs) in the 3´UTRs of platelet and reticulocyte mRNAs. Our work reveals a surprising consequence of tissue-specific regulation of core cellular machinery and provides insight into how multicellular organisms achieve tissue-specificity using a common set of biological pathways.

Results

Rescue/recycling factors are down regulated during differentiation

To test whether the accumulation of 3´UTR ribosomes is linked to regulation of ribosome termination or recycling, we surveyed the relative protein expression levels of factors involved in termination and recycling. Interestingly, PELO (and its cofactor HBS1L) are differentially expressed in mononuclear cells and platelets/reticulocytes (Fig. 1a), while levels of ABCE1 are comparable among the primary cell types tested. Consistent with these results, hemin-induced erythroid differentiation of K562 cells led to a spike in
PELO/HBS1L levels, followed by their subsequent decline after ~48 hours (Fig. 1b, Fig. 7a) and a decline in ABCE1 levels (Fig. 1b). Notably, other proteins involved in translation such as eIF1A or eRF1 did not change during differentiation, nor did the abundance of ribosomal proteins themselves (Fig. 1b). Meg01-derived PLPs lack detectable PELO expression, and have reduced ABCE1 levels, while Meg01 cells express both proteins robustly (Fig. 1c). PLPs isolated as early as 1 hour after release by Meg01 cells failed to show detectable PELO expression, indicating the PELO protein is likely excluded from nascent PLPs despite its presence in Meg01 cells (Fig. 1d). ABCE1 levels are stably reduced (but not absent) in PLPs, perhaps reflecting its essential role in eRF1-mediated hydrolysis of peptidyl-tRNA during translation termination. Taken together, these data demonstrate while ABCE1 levels decline early in differentiation, PELO/HBS1L levels are first upregulated and remain high for ~4 days (peaking at ~5-fold above levels in undifferentiated cells) before declining late in differentiation (days 5 and beyond). Thus, we hypothesized that the loss of both ABCE1 and PELO/HBS1L is responsible for the dramatic accumulation of 3’UTR ribosomes in platelets and reticulocytes.

**Selective loss of cytoplasmic protein synthesis by PLPs**

Ribosomes, assembled in the megakaryocyte, must be delivered to nascent proplatelets along with other contents designated for transport into platelets. Once formed, however, platelets have no means to regenerate ribosomes. 18S and 28S ribosomal RNAs, as well as the specific ribosomal protein RSP6, are present in young PLPs (Fig. 2a-d), and (like reticulocytes) cytoplasmic rRNA is degraded during ~24 hours after
release. This is consistent with the relative depletion of ribosomal proteins in platelets relative to other tissues (See Chapter II). Similar to reticulocyte maturation, this suggests that young platelets are most translationally active; and which, after final maturation in circulation, become translationally arrested. Unlike reticulocytes that have degraded their mitochondria\textsuperscript{24}, platelets maintain functioning mitochondria throughout their lifespan as they rely on oxidative phosphorylation for energy production\textsuperscript{25}.

Using 35S-methionine incorporation to examine global protein synthesis in Meg01 cells and PLPs we found that, after normalizing for ribosome abundance, PLPs collected between 0 and 24 hours after release were only modestly less translationally active than Meg01 cells (Fig.\textit{2e-f}). Moreover, using a ribosome transit-time assay, we calculated the average time spent by the ribosome decoding an mRNA in PLPs (53 +/- 5 seconds) (Fig. 2g), and found it comparable to other nucleated cell/tissue types such as hamster ovary cells (23 seconds)\textsuperscript{26}, placenta (56 seconds)\textsuperscript{27}, muscle cells (30 seconds)\textsuperscript{28}, or liver (30 seconds)\textsuperscript{29}. This indicates that the protein synthesis machinery in PLPs (and presumably in platelets) is fully intact and operational at the time of their release by Meg01 cells. Thus, the selective arrest of cytoplasmic protein synthesis by platelets in circulation parallels the maturation process undergone by reticulocytes.

\textbf{PELO/HBS1L induction compensates for loss of ABCE1}

ABCE1 is the principle factor involved in ribosome recycling in mammalian cells\textsuperscript{30-34}. After hemin induction, \textit{PELO/HBS1L} are induced, peaking at ~4.5 fold above their levels in uninduced cells at around 48 hours, while ABCE1 levels begin to drop immediately (Fig. 1b and Fig. 9a). To further understand the mechanistic basis for this 3’UTR
ribosome accumulation, we generated Meg01 and K562 cell transgenic cell lines that
constitutively overexpress PELO/HBS1L (Fig. 3a, Fig. 5a) and transiently knocked
down ABCE1 in these cells (Fig. 8b). As a control we used cells which transgenically
express GFP. We found, indeed, that ABCE1 knockdown leads to the accumulation of
unrecycled ribosomes in the 3’UTR as well as paused ribosomes on the stop codon,
consistent with an essential role for ABCE1 in promoting both termination and recycling
(Fig. 8c). Both of these features were rescued by overexpression of PELO/HBS1L,
indicating that the loss of ABCE1 leads to a failure in ribosome recycling that can be
rescued by PELO/HBS1L (Fig. 8c, pink traces). Thus, the initial cause of 3’UTR
ribosome accumulation is likely loss of ABCE1 and endogenous upregulation of
PELO/HBS1L can functionally compensate for loss of ABCE1 during the early phase of
differentiation.

The initial cause of the reduction in ABCE1 levels during differentiation is
unknown, but could be a consequence of the unique mitochondrial degradation process
that occurs during erythroid development (mature red blood cells lack mitochondria and
rely solely on anaerobic metabolism; See Discussion). Importantly, the mitochondria is
an important site of FeS cluster biogenesis in mammalian cells and ABCE1 requires a
FeS cluster prosthetic group for activity (Fig. 8e). Consistent with this possibility,
mitochondrially-encoded RNA transcripts were strikingly reduced in K562 cells after 2
days of hemin-induced differentiation compared to nuclear encoded transcripts (Fig. 8d).

**Transgenic restoration of PELO/HBS1L clears 3’UTR ribosomes**
We prepared ribosome profiling libraries from GFP or PELO/HBS1L-overexpressing Meg01 cells and PLPs. This led to a dramatic reduction in 3´UTR ribosomes in PLPs (Fig. 3b-d) and differentiated K562 cells (Fig. 5c), confirming that 3´UTR ribosomes are 

*bona fide* PELO/HBS1L substrates in both erythroid and megakaryoid cell lineages. Interestingly, at day 0, when there is little no elevation in 3´UTR occupancy, overexpression of PELO/HBS1L has no effect on translation. At days 2 and 4, 3´UTR ribosomes begin to accumulate (Fig. 5b-c); at this time, 3´UTR ribosomes were not significantly reduced by over expression of PELO/HBS1L (Fig. 9a), indicating that ribosomes at this state may still be associated with eRF1, or another factor, that prevents access to the ribosomal A site\(^{35-37}\). Together, these data indicate that the initial accumulation of 3´UTR ribosomes in differentiating K562 cells is caused by loss of ABCE1 and compensated by upregulation of PELO/HBS1L. The eventual loss of both ABCE1 and PELO/HBS1L likely leads to the dramatic accumulation of 3´UTR ribosomes observed in our initial primary platelet and reticulocyte ribosome profiling experiments.

**Ribosome rescue by PELO/HBS1L promotes hemoglobin synthesis**

The induction of PELO/HBS1L at the onset of differentiation in K562 cells was surprising and mirrored the kinetic rise in hemoglobin synthesis (Fig. 1b, Fig. 9a). Moreover, at days 0-2 of K562 cell differentiation, the accumulation of fetal hemoglobin protein is driven by a strong increase in fetal hemoglobin translational efficiency (TE; ribosome footprint density divided by the RNA abundance) of >70-fold for alpha globin mRNA and >7-fold for gamma globin mRNA (Fig. 7a-b, Fig. 9b). We wondered
whether the translational stimulation of hemoglobin mRNA was due in part to an increase in free ribosomes released by high expression of PELO/HBS1L. To test this possibility, we silenced PELO/HBS1L with an HBS1L-targeted shRNA (Fig. 9c, 9f) and measured fetal hemoglobin expression after 48 hours by western blot (Fig. 9g) and flow cytometry (Fig. 9d). PELO/HBS1L silencing indeed led to a reduction in fetal hemoglobin expression, but not other proteins such as eIF1A (Fig. 9g) or Actin (Fig. 9c) suggesting that induction of endogenous PELO/HBS1L plays a role in the expression of fetal hemoglobin. Consistent with this possibility, a higher percentage of fetal hemoglobin expressing cells were observed after hemin-induction at 48 hours when PELO/HBS1L was overexpressed versus GFP control cells (Fig 9e). Thus, PELO/HBS1L promotes hemoglobin synthesis, likely through increasing the available pool of ribosomes which, because of the high initiation rate of globin chain synthesis, preferentially favors the production of hemoglobin.

To confirm this possibility, we depleted ribosomes by knocking down RPS19, a ribosomal protein gene commonly mutated in Diamond Blackfan Anemia (DBA)\(^ {38} \). We induced RPS19 knockdown K562 cells with hemin and monitored fetal hemoglobin synthesis 48 hours after induction. Fetal hemoglobin synthesis was significantly impaired in cells with reduced RPS19; and strikingly, over expression of PELO/HBS1L restored fetal hemoglobin synthesis to near wild type levels (Fig 9h). This suggests that in the early phase of hemin-induced differentiation, hemoglobin translation is limited by ribosome availability, and the induction of endogenous PELO/HBS1L favors the production of hemoglobin by increasing the availability of free ribosomes. Notably, the observation that free ribosomes (having access to any mRNA) would become
disproportionately engaged in the translation of specific transcripts (like hemoglobin) is a key prediction of early models of protein synthesis\textsuperscript{39-41} and stems from the high initiation rates of globin mRNAs. Notably, this is consistent with the longstanding observation that hemoglobin mRNA has a higher affinity for the 43S preinitiation complexes that other mRNAs\textsuperscript{42}.

**Discussion**

Previous work in yeast and insect cells has implicated Dom34 and Hbs1 as interacting ribosome “rescue” factors involved in mRNA quality control pathways which degrade aberrant mRNAs\textsuperscript{14,16,43-45}. These cellular mRNA surveillance systems recognize and dissociate elongating ribosomes\textsuperscript{12,14,15,32} whose translocation is impeded by stable RNA structure\textsuperscript{43}, rare codons\textsuperscript{43}, or polybasic sequences\textsuperscript{46}, targeting aberrant, damaged, or defective mRNAs for degradation by endonucleolytic cleavage\textsuperscript{43} followed by 5’ to 3’ or 3’ to 5’ exonucleolytic mRNA decay\textsuperscript{47}. Additional work *in vitro*\textsuperscript{32} and *in* yeast\textsuperscript{18} has shown that Dom34/Hbs1 also targets vacant 80S ribosomes lacking peptidyl-tRNA such as are produced after failure of normal termination and recycling by ABCE1. Still, physiologic roles for Dom34/Hbs1 homologues in higher eukaryotes are largely unknown, though a role in relieving ribosome stalling in mouse brain has been reported by Ishimura *et al*\textsuperscript{48}, but only in mice lacking a specific tRNA.

We have shown that PELO/HBS1L rescue vacant unrecycled ribosomes in human cells and are specifically regulated during platelet/red blood cell differentiation. Our results demonstrate a switch in the ribosome recycling machinery used by cells during erythroid (and perhaps megakaryoid differentiation) – from primarily ABCE1-mediated
recycling to PELO/HBS1L-mediated recycling – before the loss of ribosome recycling factors. Additionally, we identified a key consequences of this switch as, in part, biasing differentiating K562 cells toward to the translation of fetal hemoglobin mRNA, a key step in the differentiation of red blood cells. Therefore, these data suggest that the differentiation-dependent switch in ribosome recycling factors we identified may play a important role in proper red blood cell maturation. Finally, this work also suggests a potential new avenue for therapeutic reactivation of fetal hemoglobin expression, which is normally silenced after fetal development, but continued expression has been proven a valuable therapeutic modality for disorders affecting adult hemoglobin such as sickle cell disease and beta-thalassemias.

**Mechanism of ABCE1 silencing in K562 cells**

Rapidly growing yeast (and likely mammalian cells as well) rely on ABCE1 (Rli in yeast) for termination and recycling of ribosomes at the end of the translation cycle. ABCE1 is an unusual ATP-binding cassette (ABC) transporter family member that lacks a transmembrane segment, making it unlikely to have a role in transport. Moreover, ABCE1 has a highly conserved N-terminal region consisting of eight cysteine residues that coordinate two [4Fe-4S]^{2+} clusters. Seven of the eight cysteine residues are essential for viability (at least under aerobic conditions) and the ribosome recycling function of ABCE1. Interestingly, the iron-sulphur cluster domain of ABCE1 is the only site of direct contact between ABCE1 and PELO, which binds to the empty ribosomal A site. Binding of ABCE1 to PELO in the ribosomal A site stabilizes a conformational rearrangement of the C-terminal domain (CTD) of PELO that reaches
through the A site and contacts the P site tRNA, reminiscent of the movements that
would be taken by the GGQ motif (absent in PELO) of eukaryotic release factor 1 (eRF1)
during peptide hydrolysis\(^5\). It has been proposed the FeS clusters coordinated by
ABCE1, and required for ABCE1-mediated ribosome recycling, account for the essential
requirement of mitochondria for viability of eukaryotes, as the major site of FeS cluster
biogenesis\(^{58-60}\). Since mature red blood cells lack mitochondria\(^{24,61-69}\), it is tempting to
speculate that the loss of FeS cluster biogenesis in maturing red blood cell leads to the
inactivation of ABCE1-mediated recycling observed in our study. Consistent with this
possibility, we found that hemin induction led to a striking loss of global mitochondrial
transcript abundance (Fig. 8d,e). Upregulation of PELO/HBS1L at the onset of
differentiation thus allows cells to maintain adequate ribosome recycling efficiency with
reduced levels of ABCE1. Thus could occur either because ribosome rescue under these
conditions proceeds independently of ABCE1\(^\text{12}\) or because PELO/HBS1L levels are
normally limiting for ribosome rescue by the PELO/HBS1L/ABCE1 complex. In the
latter case, upregulation of PELO/HBS1L could increase the recruitment of remaining
ABCE1 to stalled ribosomes.

**Gene-specific regulation by general translation factors**

The first gene responsible for Diamond-Blackfan Anemia (DBA\(^4\))\(^79\) was identified in
1999 as *RPS19*, which encodes a ribosomal protein of the small subunit\(^38\). How a
mutation\(^5\) encoding a protein involved in the function of a ubiquitous structure such as
the ribosome could cause a tissue-specific phenotype ignited major research and clinical

---

\(^4\) Louis Diamond and Kenneth Blackfan are the names of the pediatricians who first described the
syndromic presentation.

\(^5\) In DBA, mutations in the small or large ribosomal subunit are haploinsufficient.
interest, leading to the identification of at least 13 novel “ribosomopathies” caused by mutations in at least 26 difference genes. The mechanism(s) of tissue-specificity is unknown in all cases, but functional specialization of ribosomal proteins or ribosomes themselves has been proposed. In RPS19 knockdown cells, ribosome assembly is blocked and therefore, nucleolar dysfunction, a shortage of ribosomes, or both have been proposed as causing the reduced numbers of erythroid progenitors that causes to DBA. Here, our results uncovered a role for PELO/HBS1L in regulating ribosome availability during hemin-induced differentiation of K562 cells. The importance of this observation is underscored by the requirement for endogenous up regulation of PELO/HBS1L to support hemoglobin synthesis as well as the ability of over expression of PELO/HBS1L to functionally rescue the defect in hemoglobin synthesis that results from RPS19 depletion. Remarkably, this indicates that gene-specific changes caused by loss of a specific ribosomal protein can be rescued by overexpression of generally-acting ribosome recycling factors, which presumably increase the pool of free ribosomes. Thus, the tissue specificity observed from ribosomal protein haploinsufficiency likely results from the high rate of cell division, and thus demand for ribosomes, required by erythroid progenitors. Reduced ribosome availability has been noted previously to impair the translation of important erythroid mRNAs such as GATA1.

Notably, our results are in agreement with quantitative models of protein synthesis by reticulocytes which predict that a small change in the availability of free ribosomes – which in our case we estimate to be between 1 and 10% of total cellular ribosomes – will disproportionately result in the increased translation of mRNAs with high rates of initiation such as those encoding hemoglobin. Thus, we have
described a framework consistent with both our own data and kinetic models of protein synthesis which explain how a change in general ribosome recycling factors leads to preferential changes in the translation of individual genes.

Mechanism of PELO/HBS1L induction in K562 cells

During red blood cell maturation, the HBS1L-MYB locus undergoes progressive silencing, coordinated by direct binding of erythroid transcription factors that leads to reduced HBS1L and MYB mRNA levels. Since HBS1L and PELO are regulated coordinately in cells (Fig. 9c, 9f, 9i), reduced transcription from the HSB1L-MYB locus would be predicted to also reduce PELO levels during erythroid differentiation. DNA sequence polymorphisms in the intergenic enhancer region of the HBS1L-MYB locus give rise to common alleles that are associated with persistent fetal hemoglobin (HbF) expression throughout life (called Hereditary Persistence of Fetal Hemoglobin, or HPFH), a trait which, remarkably, ameliorates anemia caused by hemoglobinopathies like sickle cell disease (SCD) and ß-thalassemias. The mechanism through which HPFH alleles lead to increased HbF expression throughout life is poorly understood but has focused on the role of MYB, an essential erythroid transcription factor dispensable for embryonic erythropoiesis but required for adult erythropoiesis. Surprisingly, the effects of HPFH alleles on HBS1L and MYB levels have not yet been reliably defined through experiments, but an inverse correlation between HbF expression and MYB levels in erythroid precursors has been observed. Our results support a mechanism whereby higher levels of PELO/HBS1L promote HbF expression through increased translation of HbF mRNA.
Role of PELO/HBS1L regulation of mRNA stability/decay

Developing erythroid cells undergo a period of highly efficient ribosome and mRNA production before transcription ultimately arrests and pyknotic nuclei are extruded. In our data, overexpression of PELO/HBS1L led to significant reductions in 3′UTR ribosomes only in the late-phase of K562 differentiation (day 6). Notably, the acquisition of sensitivity to PELO/HBS1L expression coincides with the natural loss of PELO/HBS1L levels that we observed to occur between days 4-6 (Fig. 1b). As unrecycled 3′UTR ribosomes rapidly accumulate during days 2-4 (as opposed to days 0-2 or 4-6), but are not rescued by PELO/HBS1L, nor is mRNA decay detectably increased at these times, one explanation is that the ribosomal A site of unrecycled ribosomes is still occupied by eRF1 from termination, or another A site binding factor that prevents access of PELO/HBS1L and prevents PELO/HBS1L-dependent decay. Thus, unrecycled post-termination ribosomes that initially accumulate in a PELO/HBS1L-resistant configuration (perhaps because of continued eRF1 association) gradually become sensitive to PELO/HBS1L through a process that could involve the eEF2-mediated release of eRF1. Messenger RNA sensitivity to PELO/HBS1L-dependent decay is therefore, expected to correlate with the recognition and release of vacant 3′UTR ribosomes, perhaps via the defined role for PELO/HBS1L in recruiting RNA endonuclease(s) that stimulates mRNA decay. Consistent with this expectation, rescue of vacant ribosomes by Dom34 in yeast is genetically linked to the degradation of the mRNA as vacant 3′UTR ribosomes are enriched in yeast lacking Dom34 and Ski. Thus, during the late phase of K562 differentiation, loss of PELO/HBS1L may promote
increased mRNA stability by reducing post-termination ribosome recycling, though further experiments are needed to fully elucidate this relationship.

Materials and Methods

Generation of stable human cell lines For PELO/HBS1L overexpressing Meg01 and K562 cell lines, naïve cells were grown as described previously (See Chapter 2 Methods) and maintained between 5 and 30 passages. Cells were transduced with lentivirus and assayed for mCherry or EGFP expression between two and five days later. Positive cells were sorted and quarantined in PenStrep (Gibco) for 5-7 days, after which no antibiotics were used in the media. Cells were monitored for stable expression by flow cytometry and western blotting using PELO (abcam) or HBS1L (protein-tech) specific primary antibodies.

FACs and flow cytometry of human cells lines Cell sorting was performed using a FACsAria IIu (Becton Dickinson) using excitation lasers 488nm (for GFP) and 633nm (for mCherry). For analytical flow cytometry, the Guava EasyCyte (Millipore) instrument was used with excitation at either 488 nm (for GFP) or 532nm (for mCherry).

Lentivirus production Plasmids pPLP1, pPLP2, pVSVG were provided by S. Eacker. iDuet101a was obtained as a bacterial stab from Addgene and transformed into Stbl3 competent E. coli cells (Invitrogen). cFUGlw and cFUG were also provided by S. Eacker and transformed into Stbl3 cells. Lentiviral packaging was performed in HEK293FT cells.
(Life Technologies) using standard protocols. PELO-T2A-mCherry and HBS1L-T2A-mCherry were subcloned into cFUGlw for packaging.

**Lentiviral transduced cell lines** Meg01-GFP1, Meg01-PH1, Meg01-P1, K562-GFP3, K562-GPF4, K562-PH3, K562-PH4 were produced by transduction of indicated lentiviruses prepared as described above. Cells were transduced with virus, incubated for 2-3 days in the tissue culture incubator and then analyzed by FACs. GFP or mCherry positive cells were sorted as stable cell lines.

**shRNA sequences**

PELO: 5’-GCCAAGAAGCAGTGGGATATT-3’ (sense) and 5’-TATCCCCACTGCTTCTTGCTT-3’ (antisense)

HBS1L: 5’-GGATGAAACTGGCGAAGAATT-3 (sense) and 5’-AAGGATGAAACTGGCGAAGA-3’ (antisense)

ABCE1: 5’-GAGGAGAGTTGCAGAGATTTT 3’ (sense) and 5’-AATCTCTGCAACTCTCCTCTTT 3’ (antisense)

EXOSC8: 5’-TGGAAGAGAAGTGGGTGGAATT-3’ (sense) and 5’-TTCACCAAGTTCTCTTCCATT-3’ (antisense)

EXOSC9: 5’-GGAAGAGATGTCTAAGAAATTT-3’ (sense) and 5’-ATTTCTTAGACCATCTTCTTCTTT-3’ (antisense)

DIS3: 5’-CAGGAGAAACAAAGAGAAATT-3’ (sense) and 5’-TTTCTCTTTGTTTCTCCTGTT-3’ (antisense)
**SDS-PAGE, protein transfer** Cell samples for western blotting were prepared by lysis in polysome lysis buffer (PLB) supplemented with mini EDTA-free protease inhibitor cocktail (Roche). Cell pellets were washed twice in sterile, room temperature PBS, then mixed with an equal volume of PLB, vortexed for 10-30 seconds, and incubated on ice for 10 minutes. Lysates were clarified at 20,000 x g at 4°C for 10 minutes. Supernatants were transferred to clean microfuge tubes and stored at -80°C. Total protein concentration was determined using either the BCA (Pierce) or EZQ assay (Life Technologies). Two to twenty micrograms of total protein were combined with 5X SDS buffer and heated for 80°C for 3 minutes in a thermomixer then transferred to ice. 4-12% BisTris precast gels (BioRad Criterion) were used for SDS-PAGE and gels were run at 15-20 mW constant power until the loading dye front reached the bottom of the gel. Proteins were transferred to PVDF membrane using the TurboTransfer apparatus (BioRad) mixed molecular weight settings according to the manufacturer’s protocol.

**Western blotting** PVDF membranes were blocked immediately post-transfer in 50 ml of 5% non-fat dry milk in TBS (or PBS) with 0.1% Tween-20 for 1 hour at room temperature on an orbital shaker. After blocking, PVDF membranes were vacuum-sealed in plastic bags with primary antibody diluted in 5-10ml of blocking buffer and incubated at 4°C (cold room) fixed to the nutator overnight (8-12 hours). Membranes were washed in 35ml of TBST (or PBST) for 3 cycles of 10 minutes at room temperature on the orbital shaker. 25ml of secondary HRP-conjugated antibody diluted 1:10,000 was added and incubated at room temperature for 30 minutes to 1 hour on the orbital shaker. Finally, membranes were washed for 6 cycles of 5 minutes at room temperature in TBST (or
PBST) prior to chemiluminscent detection using femto substrate diluted 1:10 in pico substrate (Pierce). Images were developed using a film processor (Kodak). When indicated, 5% BSA was used during blocking and all antibody incubations.

**Image processing** Western blot images were scanned from film using a laser scanner (HP Intel) and analyzed using ImageJ64 software (NIH). When indicated brightness/contrast settings were adjusted, but no other changes to the images were made. For quantification, protein bands of interest were highlighted using the rectangular selection tool and Integrated density values were taken using the *Analyze>Measurements>Set Measurements* sequence.

**EZQ dot blotting and analysis** Protein quantitation was performed according to manufacturer’s protocol. Dlot blots were scanned using the Typhoon instruments using a 488nm laser for excitation and long pass filter (LPF) to collect the emission. Images were analyzed in ImageJ64 software (NIH) as follows. Images were inverted using the *Edit>Invert* option, background was subtracted automatically using the *Process>Subtract Background* option, and integrated density (IntDen) measurements were taken using the *Analyze>Measurements>Set Measurements* sequence.

**Figure List**

Figure 1. Ribosome rescue/recycling factors are regulated during differentiation

Figure 2. Translational arrest by PLPs

Figure 3. Restoration of *PELO/HBS1L* releases 3’UTR ribosomes in PLPs
Figure 4. PLP gene models

Figure 5. Restoration of PELO/HBS1L in differentiated K562 cells

Figure 6. Clearance of 3’UTR ribosomes is associated with mRNA decay

Figure 7. Translation, not RNA abundance, drives hemoglobin accumulation in K562 cells

Figure 8. ABCE1 silencing leads to ribosome recycling failure and is rescued by PELO/HBS1L over expression

Figure 9. PELO/HBS1L induction promotes fetal hemoglobin synthesis by increasing free ribosomes

Figure 10. Model – A differentiation-dependent switch in ribosome recycling factors favor hemoglobin translation

Figure S1. Purification of transgenic PLPs

Figure S2. Evidence for hyperphosphorylated forms of HBS1L

Figure S3. PELO/HBS1L overexpression interferes with normal XBPI translation

**Figure Legends**

**Figure 1. Ribosome rescue/recycling factors are regulated during differentiation** | a. Western blots for PELO, HBS1L, and ABCE1 from primary human peripheral blood mononuclear cells (PBMCs), platelets, or CD71+ reticulocytes and normalized by total protein amounts (dot blot; bottom panel). b. PELO levels spike at 48 hours and then fall to levels below those in uninduced cells around day 6. ABCE1 levels begin to decline immediately following hemin induction of K562 cells. HBG2 (fetal hemoglobin) was
monitored to ensure productive differentiation. Uninduced cells (day 0) express very little HBG2, which is strongly induced by day 2. c. Meg01-derived platelet-like particles (PLPs) lack detectable PELO levels and have strongly reduced ABCE1 levels. d. PLPs collected during the first several hours after release also lack detectable PELO expression indicating PELO is likely excluded from nascent PLPs rather than being rapidly degraded.

**Figure 2. Translational arrest of platelet-like particles (PLPs)** | a-b. PLP 18S and 28S rRNAs begin to be degraded in the first 24 hours of PLP release. c. quantification of selective loss of cytoplasmic rRNAs (28S: blue; 18S: pink), but not mitochondrial rRNA (16S: black). 18 and 28S rRNAs were quantified using the BioAnalyzer (left axis), while the mitochondrial 16S rRNA required quantitation by qRT-PCR (right axis). d. The ribosomal protein RPS6 begins to be degraded shortly after PLP release. e. 35S-Methionine incorporation by Meg01 cells and PLPs (collected from 0-24 hours after release) without (e) and with (f) normalization for ribosome abundance. g. The average ribosome transit time measured for PLPs (~53 seconds) is similar to transit times measures in other tissues and cell types.

**Figure 3. Restoration of PELO/HBS1L releases 3’UTR ribosomes in PLPs** | a. Western blot confirmed overexpression of PELO in Meg01 cells (left) and PLPs (right). b. ECDF plots of relative 3’UTR ribosome occupancy for Meg01 cells overexpressing GFP (dark blue) or PELO/HBS1L (dark red), and PLPs overexpressing GFP (light blue) or PELO/HBS1L (pink). c. RPF size distributions for transgenic PLPs. d. Meta-gene
analysis of RPFs near the stop codon. e. RPF size distributions for transgenic Meg01 cells.

**Figure 4. PLP gene models** | a. mRNA-specific ribosome profiles of PLP mRNAs demonstrating PELO/HBS1L-dependent 3′UTR ribosome release. 3′UTRs are highlighted in yellow boxes.

**Figure 5. Restoration of PELO/HBS1L in differentiated K562 cells** | a. Western blotting confirmed over expression of PELO/HBS1L in K562 cells. b. ECDF plots of relative 3′UTR ribosome occupancy from hemin-induced K562 cells overexpressing PELO/HBS1L (PH) or GFP 0, 2, or 4 days after induction. c. ECDF plots from day 6 post-hemin induction. d. Scatter plots of RPF density on coding sequences from PELO/HBS1L overexpressing cells or GFP control cells during hemin differentiation days 0, 2, and 6.

**Figure 6. Clearance of 3′UTR ribosomes is associated with mRNA decay** | a. RPF scatter plots of PELO/HBS1L overexpressing cells (PH) or GFP control cells during days 0, 2, or 6 post-hemin induction. Hemoglobin mRNAs are highlighted. b. RNA scatter plots of PELO/HBS1L overexpressing cells (PH) or GFP control cells during days 0, 2, or 6 post-hemin induction. Firefly luciferase (FLUC) *in vitro* transcribed RNA was spiked-in to each sample (at an amount predicted to give 0.1% of reads) as an external control for overall RNA content in each sample. The % of total mapped reads which mapped to the FLUC transcript are reported. c. PLPs overexpressing PELO/HBS1L or
GFP control PLPs were collected over 7 days and analyzed by qRT-PCR for the relative abundance of specific mRNAs using the mitochondrial transcript COX1 as a control (blue bars; fold change in RNA level. 3’UTR clearance scores (CS) are also indicated (gray bars), where CS = relative 3’UTR ribosome occupancy[GFP sample] / relative 3’UTR ribosome occupancy[PELO/HBS1L sample]. A significant inverse correlation was observed between change in RNA level and release of 3’UTR ribosomes by PELO/HBS1L.

**Figure 7. Translation, not RNA abundance, drives hemoglobin accumulation in K562 cells** | 3’UTR ribosome occupancy was plotted against translational efficiency (TE), where the TE for each mRNA is calculated according to TE=CDS (RPF reads) / CDS (RNA reads). Hemoglobin mRNAs are highlighted. a-b. Between days 0 and 2, hemoglobin mRNAs undergo a dramatic (>30 fold) increase in translational efficiency. This increase in translational efficiency is fully reversed by day 6 when the TE of hemoglobin mRNAs fall to just below their levels in uninduced cells (c). At day 6, 3’UTR ribosomes are modestly reduced by PELO/HBS1L over expression, fewer mRNAs are detected overall, and a group of well-translated mRNAs at day 6 in GFP control samples are either absent or poorly translated when PELO/HBS1L levels are restored. d. Total protein synthesis was monitored by 35S-methionine incorporation.

**Figure 8. ABCE1 silencing leads to ribosome recycling failure and is rescued by PELO/HBS1L over expression** | a. Experimental design for ABCE1 knockdown by shRNA. b. Western blotting confirmed knockdown of ABCE1 at 2 days post-transfection in K562 cells overexpression PELO/HBS1L (pink box) or GFP control cells (purple box).
c. Ribosome profiling of ABCE1 knockdown cells in the presence of transgenically overexpressed PELO/HBS1L (pink) or GFP control (purple). Meta-gene analysis was performed for RPFs around the stop codon to look for ribosome recycling failure. Transfection with a control shRNA (left) did not lead to an accumulation of 3’ UTR ribosomes, knockdown of ABCE1 in GFP control cells (purple) led to a substantial increase in terminating ribosomes paused on the stop codon (black arrow; upper) as well as the accumulation of unrecycled 3’ UTR ribosomes (black arrow; lower). Both of these features were not observed in ABCE1 knockdown cells that overexpress PELO/HBS1L (pink traces). d. Hemin induction of K562 cells led to a major reduction in transcripts encoded by the mitochondrial genome (red) relative to nuclear-encoded transcripts (gray), consistent with the mitochondrial breakdown that occurs during erythroid differentiation. e. The ABCE1 protein has an N-terminal FeS cluster coordination center for $[4\text{Fe}4\text{S}]^{2+}$ that is required for ribosome recycling activity$^{56}$ and forms direct contacts with PELO$^{94}$. FeS clusters are generated in the mitochondria, perhaps explaining the loss of ABCE1 from differentiating erythroid cells.

**Figure 9. PELO/HBS1L induction promotes fetal hemoglobin synthesis by**

**increasing free ribosomes | a.** Between 6 and 12 hours after hemin induction, PELO/HBS1L are induced in K562 cells and rise with similar kinetics as the rise in hemoglobin (Hb gamma). b. Increase in translational efficiency (TE; ribosome occupancy in the CDS normalized to mRNA abundance) of fetal hemoglobin mRNAs during K562 cell differentiation. Alpha and gamma globin genes encode the two proteins making up the fetal hemoglobin (HbF) heterodimer. c. Modest reduction in
PELO/HBS1L using a shRNA against HBS1L leads to reduced fetal hemoglobin accumulation. Notably, knockdown of HBS1L is sufficient to induce a commensurate reduction in PELO levels in multiple cell types, for example HEK293 cells (f). d. Flow cytometry analysis of fetal hemoglobin expression after hemin-induction also revealed that knockdown of PELO/HBS1L reduces the accumulation of fetal hemoglobin. Transfection of a shRNA against PELO alone did not produce effective knockdown of the PELO protein. e. Overexpression of PELO/HBS1L modestly increases the percent of fetal hemoglobin positive cells at 48 hours after hemin induction. g. Knockdown of PELO/HBS1L using shRNAs against HBS1L and PELO abrogates fetal hemoglobin induction. Notably, further knockdown of ABCE1 does not further reduce hemoglobin production h. knockdown of RPS19, a cause of reduced ribosome availability implicated in Diamond Blackfan Anemia (DBA), leads to a reduction in fetal hemoglobin levels. The abrogation of hemoglobin synthesis could be rescued by overexpression of PELO/HBS1L (PH in figure). i. knockdown of the exosome core component EXOSC8 led to an increase in PELO/HBS1L levels and prevented fetal hemoglobin synthesis.

Figure 10. Model – A differentiation-dependent switch in ribosome recycling factors favors hemoglobin translation | a. Schematic representation of the “switch” in ribosome recycling machinery in differentiating K562 cells as well as the accumulation of unrecycled ribosomes. b. Similar representation of the changes in fetal hemoglobin RNA (red) and protein (light blue). Total protein synthesis (measured by 35S-Methionine incorporation as shown in Figure 7d is also shown. c. Model for a differentiation-dependent switch in ribosome recycling factors favoring hemoglobin synthesis.
**Figure S1. Purification of transgenic PLPs** | a. PLPs were purified from Meg01 cells stably expressing GFP (b) or PELO/HBS1L-T2A-mCherry (c). d. Quantitation of fluorescence signal from PLPs.

**Figure S2. Evidence for hyper-phosphorylated forms of HBS1L** | a. High molecular weight forms of HBS1L were detected in HEK293 cells, K562 cells, and primary reticulocytes. To determine whether phosphate groups contribute to the aberrant migration of HBS1L during SDS-PAGE, we treated lysates with alkaline phosphatase (Alk Phos). Modest size reductions could be appreciated in treated lysates suggesting a contribution of hyper-phosphorylation to the aberrant migration pattern of HBS1L.

**Figure S3. Long-term G418 selection upregulates PELO/HBS1L and reduces 3´UTR ribosomes in PLPs** | a. ECDF plots of relative 3´UTR ribosome occupancy in transgenic PLPs derived from Meg01 cells selected in G418 (an aminoglycoside)-supplemented culture media. b. Western blot showing upregulation of PELO/HBS1L by Meg01 cells and PLPs after long-term growth in G418.

**Figure S4. PELO/HBS1L overexpression interferes with normal XBP1 translation** | a. IGV browser views of RNA-Seq (top two panels) or ribosome profiling data (bottom two panels) from PLPs overexpressing GFP or PELO/HBS1L focused on the XBP1 locus. The mRNA segment remove by IRE1-dependent cytoplasmic splicing at the ER membrane is indicated in light blue. The spliced (XBP1s) and unspliced (XBP1u)
transcripts are indicated. **b.** XBP1-specific ribosome profiles from Meg01 cells (left panels) and PLPs (right panels). **c.** Ribosome profiles from rabbit reticulocyte lysate *in vitro* translation reactions programmed with an *in vitro* transcribed XBP1 C-terminus containing firefly luciferase RNA (top) or control luciferase RNA (bottom).
References


29. Gehrke, L., Bast, R. E. & Ilan, J. An analysis of rates of polypeptide chain elongation in avian liver explants following in vivo estrogen treatment. II. Determination of the specific rates of


Fig. 1

(a) Primary human cells

(b) K562 cells

(c) Meg01 cells

(d) PLPs (h)

- PELO
- HBS1L
- ABCE1
- Total Protein

- HB02
- PELO
- ABCE1
- aIF1
- eIF1A
- RPS6
- RPL4

- PELO
- ABCE1
- Actin

- PELO
- ABCE1
- Actin

- PELO
- ABCE1
- Actin

- PELO
- ABCE1
- Actin
Fig. 5

a

K562 cells

GFP PH

PELO

HBS1L
eRF1

b

0d 2-4d

GFP 0d

PH 0d

GFP 2d

PH 2d

GFP 4d

PH 4d

Relative 3' UTR occupancy [log2]

ECDF

0.0 0.2 0.4 0.6 0.8 1.0

Relative 3' UTR occupancy [log2]

ECDF

0.0 0.2 0.4 0.6 0.8 1.0

c

6d

GFP 6d

PH 6d

d

0d

2d

GFP density (CDS)

RPF

GFP

0d 2d 6d

PELO/HBS1L

PELO/HBS1L

PELO/HBS1L

115
Fig. 6

a

RPF–day0

Hb-gamma
Hb-alpha
Hb-beta

RPF–day2

RPF–day6

GFP

b

RNA–day0

GFP 0.1% FLUC
PH 0.13% FLUC

RNA–day2

GFP 0.05% FLUC
PH 0.04% FLUC

RNA–day6

GFP 0.09% FLUC
PH 0.40% FLUC

GFP
Fig. 6 continued

C

PLPs

spearman rho = (rho=0.224)

RNA foldChange

3'UTR CS [log10]

RNA foldChange (7d)

utr3 Clearance Score [log10]
Fig. 7

a

Day 0

3'UTR occupancy

Translation Efficiency (TE)

b

Day 2

3'UTR occupancy

Translation Efficiency (TE)

c

Day 6

3'UTR occupancy

Translation Efficiency (TE)

Hemoglobin translational stimulation (~32 fold)

Hemoglobin translation is reduced
Fig. 7 continued

![Graph showing 35S cpm x 10^5 over days 0 to 10 for control and +Hemin conditions, with significance levels indicated for days 1-5 and days >5.]
Fig. 9 continued

<table>
<thead>
<tr>
<th></th>
<th>g</th>
<th>h</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>48 hrs post-knockdown</td>
<td>48 hrs post-knockdown</td>
<td>48 hrs</td>
</tr>
<tr>
<td>shCtrl</td>
<td>+</td>
<td>-</td>
<td>shCtrl</td>
</tr>
<tr>
<td>shHBS1L</td>
<td>-</td>
<td>+</td>
<td>shEXOSC8</td>
</tr>
<tr>
<td>shPELO</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>shABCE1</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>globin</td>
<td></td>
<td></td>
<td>EXOSC8</td>
</tr>
<tr>
<td>PEL0</td>
<td></td>
<td></td>
<td>HBS1L</td>
</tr>
<tr>
<td>ABCE1</td>
<td></td>
<td></td>
<td>PEL0</td>
</tr>
<tr>
<td>elf1A</td>
<td></td>
<td></td>
<td>Hb gamma</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Actin</td>
</tr>
</tbody>
</table>
Fig. 10 continued

c

- **ABCE1-mediated ribosome recycling**
  - Undifferentiated

- **PELO/HBS1L-mediated ribosome recycling**
  - Days 2-4

- **Loss of PELO/HBS1L**
  - Days 6+

- Non Hb mRNA
  - Reduced ABCE1
  - Induction of PELO/HBS1L

- mRNA "sensitized" to PELO/HBS1L-dependent mRNA decay

- Transcriptional regulation of the HBS1L-HYB locus
  - Co-regulation of PELO/HBS1L

- Continued translation of stable mRNAs

- Stimulation of Hb translation >30 fold

- Free ribosomes

- Loss of PELO/HBS1L
a HEK293 cells
HBS1L-hairpin:  - - - + + - - -
Alk Phos:       - + + + + + + +

5 sec. exposure
76 kDa · size reduction

10 sec. exposure

- Alk Phos
Chapter IV: Conclusions

In this work, we performed the first comprehensive study of translation in non-nucleated blood cells and made significant contributions to multiple diverse areas of biology. Still, many outstanding questions raised in this work remain, awaiting further experiments and analysis. Some of these are described below.

Megakaryocytes/platelets as model system for localized translation

In the megakaryocyte/platelet lineage, we documented >6,700 mRNAs that are translated by circulating platelets and found that it represents only a subset of the mRNAs available to the megakaryocyte. How are mRNAs trafficked or sorted into platelets? One possibility is that nascent platelet territories form around sites of active polysomes, which are needed to synthesize cargo proteins packaged into platelet granules. In the same way that localized translation at dendrites is critical for remodeling of synapses, continued translation by platelets may reflect localized translation by megakaryocytes. This work should serve as a starting point for understanding how mRNAs (and ribosomes) are trafficked to different locations in the cell. Is the same cellular machinery used to traffic dendritic mRNAs as in megakaryocytes? Our evidence points to a role for FMRP in this process and suggests similar machinery may used to regulate local translation in diverse tissues like brain and bone marrow. Hematopoietic zinc finger (Hzf) RNA binding protein is another promising candidate with dual roles in megakaryocyte alpha granule trafficking as well as mRNP localization in neuronal dendrites. Mice lacking Hzf have profound bleeding diathesis with lack of alpha granules in MKs and platelets.
and also impaired motor functions\textsuperscript{11} perhaps indicating a dual role in regulating mRNA localization/translation in both tissues.

\textit{mRNA stability during erythroid differentiation}

Messenger RNA stability during red blood cell maturation has been extensively studied; yet, with the advent new technology, revisiting classical biological questions will almost certainly reveal new insight. For globin mRNAs, whose intrinsic features confer exceptional stability\textsuperscript{12-14, 15-19, 20}, half lives have been measured between 30 and 60 hours. Specific 3´UTR elements are specifically involved in conferring high mRNA stability to the globin and other important erythroid genes such as transferrin and ferritin mRNAs\textsuperscript{21-24}. Alpha globin mRNA, for example, contains a cytosine-rich element (CRE) in its 3´UTR that forms a RNP complex (called the alpha complex)\textsuperscript{25-28} that physically protects the cleavage site targeted by an erythroid endoribonuclease (ErEN). When exposed to the cytoplasm by the departure of the alpha complex, the 3´UTR is endonucleolytically cleaved, leading to rapid degradation of the alpha globin mRNA\textsuperscript{29, 30}. Fascinatingly, naturally occurring mutations in the stop codon of human alpha globin genes\textsuperscript{6} lead the ribosome to translate the first 31 amino acids of the 3´UTR, which displaces the alpha complex RNP, and stimulates the mRNA’s rapid destruction by ErEN\textsuperscript{31}. This pathway is mechanistically related to non-stop decay (NSD) that has been described in yeast\textsuperscript{32, 33}, yet differs in an important way. Multiple studies have shown that, after mutating the CPE / ErEN cleavage site which renders the 3´UTR immune to alpha complex binding (or ErEN cleavage), translation into the 3´UTR still targeted the mRNA

\textsuperscript{6} These are referred to as Constant Spring (CS) mutations and are named for the place of their discovery in Constant Spring, Jamaica.
for degradation in erythroid cell types but not in several non-erythroid cell types\textsuperscript{26, 34}. As the stop codon was placed iteratively further into the 3’UTR, leading to translation of more of the 3’UTR, the resulting mRNA was progressively more unstable\textsuperscript{34}, indicating that translation of the 3’UTR in erythroid cells is particularly prone to triggering mRNA decay. In the case of constant spring mutations, elongating ribosomes, translating into the 3’UTR, could, in principle, displace vacant, slow-moving post-termination ribosomes in the 3’UTR, improving the accessibility by endoribonucleases to the 3’UTR. Consistent with this model, delta globin mRNA, which (surprisingly) accumulates far fewer 3’UTR ribosomes than the other hemoglobin mRNAs has an extremely short half life of just \(~4.5\) hours compared to half lives of 30 to 60 hours for alpha, beta, and gamma globin mRNAs\textsuperscript{35}. The stability of non-globin mRNAs is less well understood but experiments identified two distinct populations of polyA+ mRNA – one population with an average half life of 3 to 6 hours, and a second more stable population with an average half life of 30 to 36 hours\textsuperscript{15-17, 36}. Whether these groups of mRNA systematically differ in their 3’UTR occupancy by ribosomes is unknown. Additionally, hemin-induced differentiation of K562 cells promotes the stability of many non-globin mRNAs\textsuperscript{37-42}, suggesting that erythroid development leads to post-transcriptional stabilization of many mRNAs.

\textit{Regulation of ribosome recycling during hematopoiesis}

Most notably, our work uncovered that the \textit{PELO/HBS1L} genes are specifically regulated during the development of platelets and red blood cells and that this leads to the accumulation of inactive 80S ribosomes that occupy 3’UTRs. We found that cellular HBS1L levels regulate PELO levels (likely in all tissues) but that it could be especially
important during hematopoiesis when the HBS1L-MYB locus undergoes transcriptional silencing. With the advent of genome-editing technologies like CRISPR, it would now be possible to introduce specific mutations at the polymorphic sites in the HBS1L-MYB intergenic region and ask HPFH alleles modulate ribosome rescue activity by PELO/HBS1L in erythroid cells. If HPFH alleles modulate PELO/HBS1L activity by erythroblasts, what is the effect in vivo? Adoptive bone marrow reconstitution experiments could test whether erythroid progenitors derived from disease backgrounds (for example Diamond-Blackfan Anemia (DBA), sickle cell disease (SCD), or ß-thalassemia) carrying HPFH alleles are capable of interfering with disease pathogenesis.
References


Appendix I: A High-throughput strategy for profiling IRES-mediated translation

Abstract
Stress response pathways in higher eukaryotes such as nutrient deprivation, viral infection, and hypoxia lead to a global reduction in protein output. However, the translation of specific stress-related genes is often resistant, or even stimulated, by these conditions. Internal ribosome entry site (IRES)-mediated translation initiation permits mRNAs having higher intrinsic affinity for the translation initiation machinery, usually by bypassing the requirements for specific initiation factors, to maintain protein output despite a global reduction in protein synthesis. Here, we describe a novel IRES profiling strategy that relies on the translation of a covalently circular RNA reporter to render it immune to the 5’ end-dependent initiation mechanism of translation. We first prepared covalently circular RNA molecules encoding the luciferase open reading frame by a ribozyme-mediated intramolecular cyclization reaction in vitro and in HEK293 cells. We next demonstrate that circular RNAs are substrates for translation only if they contain an IRES. Finally, we adapt our circular RNA reporter system into a novel high throughput IRES profiling strategy that tests the ability of millions of unique mRNA fragments to confer translatability to the closed circular RNA reporter.

Introduction
Protein synthesis occurs in four distinct stages: initiation, elongation, termination, and recycling. Among these stages, the mechanism of translation initiation differs most
dramatically between prokaryotes and eukaryotes. However, despite different mechanisms of assembling elongation-competent 80S ribosomes, initiation results in identical base pairing of the anticodon stem loop of aminoacylated initiator tRNA with the AUG initiation codon in the P site of the ribosome in both bacteria and eukaryotes. In prokaryotes, the events leading to translation start site selection involve a Shine-Dalgarno sequence that is found ~8 nucleotides upstream of the initiation codon which directly base pairs with a complementary sequence at the 3′ end of the 16S ribosomal RNA (rRNA). This system of initiation is well suited to the translation of polycistronic messenger RNAs (mRNAs) because of the potential for translational coupling between adjacent cistrons, and requires 30S ribosomes, formylmethionine (fMet-tRNA\textsubscript{f}\textsuperscript{Met}), and three initiation factors (IF1, IF2, and IF3). In eukaryotes, mRNAs are primarily monocistronic and start codon selection occurs through a set of related 5′-end-dependent processes. First, components of the eukaryotic translation initiation machinery recognize the 7-methylguanosine cap (m\textsuperscript{7}GpppN, where N is any nucleotide) linked to cellular mRNAs co-transcriptionally via an unusual 5′-5′ triphosphate bridge. The eIF4F cap-binding complex, consisting of factors eIF4A, eIF4G, and eIF4E, assembles on m\textsuperscript{7}G-cap and facilitates the recruitment of the 43S preinitiation complex by interacting with eIF3. The 43S preinitiation complex bound to mRNA forms the 48S scanning complex, and searches along the 5′UTR in a 5′-to 3′ direction for the AUG. The interaction between the anti-codon loop of the tRNA\textsubscript{f}\textsuperscript{Met} and the AUG in the P site is aided by a specific nucleotide context gccRccAUGG (Kozak sequence) and triggers GTP hydrolysis of eIF2. This stimulates the departure of translation initiation factors. eIF5B stimulates the

---

\textsuperscript{7} In bacteria, mitochondria, and chloroplasts, the tRNA used for initiation is N-formylmethionine (fMet) while in eukaryotes the formyl- modification is not used.

\textsuperscript{8} In \textit{E. coli}, the Shine-Dalgarno sequence is AGGAGGU.
joining of the 60S ribosomal subunit, which (after GTP hydrolysis and departure of eIF5B) results in the formation of elongation-competent 80S ribosomes.

**5’ end-independent mechanisms of translation initiation by viruses**

In virally infected cells, the host translation machinery is co-opted by a single viral mRNA – in favor of thousands of other cellular mRNAs which might necessarily compete for the same initiation machinery. This is accomplished, in part, by an alternate 5’-end independent mode of translation initiation called internal ribosome entry that bypasses the requirement for many, or in some cases all, of the eukaryotic translation factors. Internal ribosome entry site (IRES)-mediated translation, thus, involves *cis*-acting RNA structures that recruit components of the translational machinery to the viral RNA stimulating translation of the viral polyprotein over other cellular mRNAs.

The detailed mechanisms involved in the recruitment of the translation machinery to viral IRESes which leads to the formation of elongation-competent ribosomes are not well understand. Examples of IRES-mediated initiation in a few specific cases, however, are broadly informative. The poliovirus IRES contains a domain that interacts with eIF4G and leads to the formation of an IRES-eIF4G-eIF4A complex that is capable of recruiting the 43S preinitiation complex. Host IRES interacting factors (ITAFs) have been identified that stimulate IRES-eIF4G binding by stabilizing a specific conformation of the IRES. Alternatively, the encephalomyocarditis virus (EMCV) IRES interacts with eIF4G and eIF4A but recruits eIF3 and 40S subunits. Consequently, both poliovirus and EMCV IRES-mediated translation initiation mechanisms bypasses the need for the cap binding protein eIF4E. In contrast, the hepatitis C virus (HVC) IRES directly
interacts with the 43S preinitiation complexe in the absence of the eIF4F cap binding complex or eIFs 1,1A, and 4B. Finally, and perhaps most strikingly, the cricket paralysis virus (CrPV) IRES is capable of directly recruiting the empty 40S subunit by folding into a tRNA-like domain that binds to the P site, molecularly mimicking the anti-codon stem loop of tRNA$_{\text{Met}}$. These differing mechanisms of translational recruitment suggest that IRES-mediated translation likely represents a range of strategies which reduce the dependence on specific host initiation factors for viral translation initiation.

**IRES-mediated initiation of cellular mRNAs**

Practical limitations to the study of alternate initiation mechanisms have hindered mechanistic understanding of host RNA structural elements that function similarly to viral IRESes$^{35-37}$. These hindrances result from the feasibility of adequately blocking 5’ end-dependent initiation pathways. The m$_7$G cap, which stimulates initiation approximately two fold *in vitro*, is not strictly required for translation$^{38}$, and uncapped RNAs are likely still translated through a 5’-end dependent initiation mechanism$^{39}$. For this reason, m$_7$G-capped and A-tailed bicistronic reporters have become the gold standard for measuring IRES activity of a RNA sequence. In this assay, cap-dependent initiation occurs on the first cistron leading to translation of a reporter protein such as firefly luciferase. The second cistron is separated from the termination codon of the first cistron by the putative IRES and encodes a distinct reporter such as renilla luciferase. When transfected into cells, or used to program *in vitro* translation reactions in lysate, the ratio of reporter proteins (e.g. RLUC/FLUC) produced by the bicistronic RNA indicates the relative contributions of cap-dependent translation (FLUC) and IRES-mediated
translation (RLUC). This approach has identified >100 cellular mRNAs with IRES activity\textsuperscript{40,41} and IRES-mediated translation has been proposed to play important roles in cell cycle regulation\textsuperscript{42}, apoptosis\textsuperscript{43}, cancer\textsuperscript{44}, and hypoxia\textsuperscript{45}.

The use of bicistronic vectors for high throughput IRES screening, however, is not feasible. First, this is because of the tendency of random RNA fragments to introduce new, cryptic splice sites into the bicistronic vector. Alternate spliced isoforms of the reporter inevitably lead to monocistronic forms of the reporter RNAs. These monocistronic forms of the reporter confound the contributions of IRES activity and reporter protein derived from monocistronic forms of the reporter. Second, the downstream cistron of a bicistronic reporter can be translated by a “leaky scanning” process that involves normal cap-dependent binding of the 43S preinitiation complex to the RNA and scanning, but instead of initiating of the first AUG encountered in the RNA, the ribosome skips it and continues scanning downstream. Thus, the ribosome can occasionally initiate on any downstream AUG in the message including the AUG of the second downstream cistron. This is problematic for the interpretation that reporter protein produced from the downstream cistron reflects internal initiation.

Here, we present an alternate strategy for high throughput profiling of IRES activity in living cells from a library of transcriptome-derived RNA fragments. Our strategy relies on the expression of a library of covalently circular luciferase RNA reporter molecules, with each one containing a library fragment upstream of the initiation codon after circularization (\textbf{Fig. 1}). This approach does not suffer from the problem of cryptic splice sites or leaky scanning because we are able to limit our analysis to only covalently circular RNA molecules. When the introduction of a cryptic splice site to the
reporter occurs, it will prevent covalent circularization and be excluded from our analysis. Leaky scanning is also not a problem for our system because any covalently circular RNA capable of binding to ribosomes must, by definition, possess internal ribosome entry activity. These ideas will be elaborated on in our results below. We first validated our strategy for preparing covalently circular RNA templates by ribozyme-mediated cis-splicing \cite{46,47}. Next we tested known cellular and viral IRESes for their ability to permit translation of the circular reporter in HEK293 cells and rabbit reticulocyte lysate. Finally, we adapted our IRES profiling assay into a high throughput strategy by preparing a library of circular RNA molecules and capturing those containing a putative IRES by deep sequencing (Fig.3).

**Results**

**Ribozyme-mediated self-splicing of circular RNA *in vitro* and in HEK293 cells**

A faithful reporter of IRES-mediated translation *in vitro* or in cells must not be subject to cap-dependent translation. This may be accomplished *in vitro* by the addition of excess m\(_7\)G cap analogues \cite{48} (that competes for available eIF4G), or in cells by treatment with hypoxia \cite{49} or expression of viral cap-cleaving proteins \cite{50}. Alternatively, cap-dependent initiation may be blocked by in the reporter alone, for example by the use of bicistronic reporter assays (described above). However, leaky scanning and cryptic splice sites have hindered their utility for high throughput IRES screening approaches. Here, we sought to use circular RNA templates for translation – a feature which, necessarily, relies on the presence of an IRES \cite{51,52} as circular RNAs are not substrates for the eukaryotic
translation machinery \(^{39}\) (Fig. 1). Thus, we generated self-splicing RNAs \textit{in vitro} and in cells using a permuted ribozyme, derived from the group I intron of T4 bacteriophage, encoded at the 5´ and 3´ ends of a luciferase open reading frame based on the strategy used by Ford and Ares \(^{46}\). We confirmed ribozyme-mediated circularization in several ways. First, we used a RT-PCR strategy with outward directed PCR primers that will only amplify a product when productive circularization has occurred (Fig. 4a-b). We quantified circularization efficiency using qRT-PCR and found that circularization occurs with an efficiency of between 80 and 100\% (Fig. 4c), with smaller insert sizes circularizing more efficiently. Next, we ran our \textit{in vitro} transcription reactions on a denaturing 6\% TBE-Urea gel and visualized the products of productive splicing reactions including linear precursor, circular product, as well as 5´ and 3´ introns that are excised from the ribozyme upon splicing (Fig. 4d). Consistent ribozyme-mediated self-splicing (rather than trans-splicing), circular products were resistant to the processive exonuclease RNaseR (Fig. 4e). To confirm circularization is dependent on an intact ribozyme, we deleted 18 nucleotides around the 5´-splice site rendering the ribozyme inactive; as predicted, this construct failed to circularize in both assays (Fig. 4a-d).

**Translational pausing on IRES-containing circRNA reporter**

When circRNAs are used to program \textit{in vitro} translation assays, or expressed in cells, productive translation events can be observed either through the production of luciferase protein (indicating a productive translational event) or the association of circRNA with ribosome complexes (indicating ribosome binding occurred) (Fig. 3, ii-iv). To adapt our system of circular luciferase reporters for the simultaneous screening a large library of
fragments for IRES activity in a single lysate or cell population, we encoded a strong nascent-chain mediated pause site derived from the human gene XBP1 gene just before the stop codon of our reporter. Thus, in order for a ribosome to stall on the pause site, it must necessarily translate the luciferase open reading frame. Using ribosome profiling, we confirmed that circRNAs containing the EMCV IRES are translated and ribosomes dwell on the specific pause site encoded by the XBP1 gene (Fig. 5a). While this strategy does not ensure that all ribosome complexes formed in association with circRNAs will have led to productive translation initiation and elongation, it does strongly prolong the dwell time of actively translating ribosomes on circular RNA and allow for efficient capture of ribosome-associated circular RNA.

Validation of a circRNA IRES reporter assay using the VEGFA 5′UTR

As a test of our circRNA IRES reporters, we prepared circRNA samples containing the predicted B-IRES contained in the human VEGFA 5′UTR or a control fragment of similar length. circRNAs were used to program in vitro translation reactions in rabbit reticulocyte lysate or transfected into HEK293 cells. After 20 minutes (for RRL) or 24 hours (for 293 cells), ribosomes, and any associated circRNAs, were pelleted over a sucrose cushion and analyzed by RT-PCR after reverse transcription using random hexamers (Fig.6). “NanoLuc” PCR primers were used to productively amplify a fragment of the reporter present in either linear or circularized RNA reporter. Splice-junction (“Spl-JCN”) primers, which lead to productive amplification across the observed spliced junction, are predicted to generate a product only from circularized RNA. As predicted circRNAs were only observed in material co-pelleted with ribosomes when an IRES was
present in the circRNA both in rabbit reticulocyte lysate as well as in HEK293 cells. To confirm circularization took place in HEK293 cells, we analyzed the presence of circularized products from both IRES-containing and control circRNAs from HEK293 cell input fractions prior to ribosome pelleting (Fig. 6c). Finally, to ensure circularization in HEK293 cells produces the same junctional spliced sequence as occurs in vitro, we TOPO-cloned junctional PCR products prepared from HEK293 input material prior to ribosome pelleting. All clones analyzed contained the predicted splice junction.

**Generation of a functional IRES profiling strategy**

To adapt our circRNA IRES reporter system to be able to screen a complex library of RNA fragments for IRES activity, we first validated that ribozyme-mediated self-splicing occurs efficiently in the context a library. To generate input libraries for IRES profiling, we began with 1 µg of human brain total RNA, choosing brain tissue because of its high diversity of expressed transcripts (Fig. S1b). We next prepared “splayed-end” strand-specific TruSeq (Illumina) libraries (Fig. 7a) that incorporate all of the Illumina adaptor sequences needed for high throughput sequencing. Using the Gibson Assembly method\(^{53}\), we generated ribozyme-reporter flanking libraries with brain RNA fragments flanked by ribozyme sequences required for self-splicing (Fig. 2, and Methods). We tested the splicing efficiency of Gibson Assembly libraries after in vitro transcription and found it to be ~70% (Fig. 7b). These libraries also maintained a high degree of complexity expressed in number of distinct molecules and covered a broad number of transcripts and transcript regions (Fig. S1b). We next transfected our Gibson Assembled dsDNA libraries into HEK293 cells along with an expression vector for T7 RNA Polymerase (T7
RNAP) (Fig. 3, i-ii). T7-dependent cytosolic transcription produces uncapped transcripts which are unstable in cells, but which are protected from cellular exonucleases by circularization. Total reporter RNA level was strongly dependent on T7 RNAP co-expression (Fig. 7c, left panel) as was luciferase activity (Fig. 7c, center panel). circRNA was detected in both input and ribosome pelleted fractions, which suggests productive splicing occurs in cells.

**IRES profiling of HEK293 cells during hypoxia**

We next prepared IRES profiling samples in HEK293 challenged with 1% O₂ for four hours to block cap-dependent translation⁴⁵ and also because hypoxia has been reported to stimulate the activity of the VEGFA IRES, our best positive control IRES. Into our starting material, we “spiked-in” eight known IRES sequences and control sequences (the antisense permutation of the known IRES) (Fig. S2). 24 hours after replicate transfections, we transferred cells to 1% O₂ while control samples were kept under ambient conditions (20% O₂) (Fig. 9a). Upon removing cells from the hypoxia (or standard) incubator, cycloheximide was added (final concentration 100 µg/ml) to the cell culture media for two minutes during processing. Ribosomes, and any associated circRNA, was pelleted by ultracentrifugation and total RNA was extracted from input and pelleted material by acid phenol/chloroform extraction. After reverse transcription using random hexamers and circle-specific PCR amplification (Fig. 3, v). Libraries were then sent for deep sequencing using a HiSeq2000. Library fragment enrichment was determined by calculating the number of unique fragments (per million reads sequenced; rpm) present in the pelleted material normalized to the number of unique fragments
present (per million) in the input material (Fig. 8a). MA plots of this information demonstrate that several spike-in IRES controls are strongly enriched by this method (VEGFA, HSPA1A, and HSPA5) as well as a large number of histone mRNAs which are known to be translated via a cap-independent (or cap-assisted) mechanism \(^{54}\). Heat shock genes, whose expression has been proposed to occur through a cap-independent mechanism, were not broadly enriched by this approach \(^{55}\). Specific gene models for VEGFA translation and IRES profiling revealed that VEGFA expression is strongly stimulated by hypoxia, and the VEGFA IRES is active in HEK293 cells and occupied by ribosomes (Fig. 8b, Fig. 9c). Other mRNAs were also strongly stimulated by hypoxia (STC1 and BHLHE40), but fragments from these genes were not present in our IRES profiling input library, rendering us unable to conclude whether their stimulation is due to IRES activity, although regions of their 5´UTRs are occupied by ribosomes. Even short durations (here 4 hours) of hypoxia leads to substantial alterations in gene expression likely at the level of translation due to the short duration of exposure including those with defined roles in the hypoxic response such as Hsp70, EGLN1, MTOR, and VEGFA (Fig. 9e).

Discussion

We have described a simple reporter assay for measuring IRES activity of any RNA fragment, for example the 5´UTR of VEGFA. This system relies on the ability of a covalently circular RNA template (circRNA) to undergo translation initiation. As circRNA is not bound by ribosomes \(^{39}\), its translation can only occur through an IRES-dependent mechanism \(^{51,52}\). While there is a single report of IRES-mediated translation
of circRNA in the literature, our system includes two major advances. First, we generated circular RNAs using a cis-acting ribozyme that, when present in a permutated orientation, releases a circular product. The approach taken previously involves a “splinted ligation” strategy whereby a specifically prepared oligonucleotide, complementary to short stretches of the 5´ and 3´ ends of the target RNA, is needed for in vitro ligation by RNA ligase. Importantly, this approach can only generate circular RNA in vitro, but not in cells. Moreover the efficiency of the intramolecular process (versus ligation of RNA molecular in trans) was not quantified. Second, the major advantage of our approach is that it allows for high throughput screening of many putative IRESes in a single experiment.

**IRES profiling of HEK293 cells**

Our IRES profiling datasets generated from HEK293 cells revealed several noteworthy features. First, our spike-in control IRES sequences, but not our negative control sequences, were strongly enriched among the ribosome-associated circRNA fragments recovered by sequencing (Fig. 8a). However, we did observe a relatively large variance for circRNA fragments detected in the ribosome-pelleted material. This is likely due to several ongoing processes including a high rate of identical sequencing reads – which may, in principle, reflect multiple copies of distinct RNA molecules with the identical sequence – but which also could reflect PCR amplification of a single RNA molecule present in multiple DNA copies during sequencing. A conservative solution to this problem is to collapse all identical sequencing reads into a single counting value, and quantitate the number of unique fragments present in the input and pelleted material. This
strategy, together with the use of paired-end sequencing – which defines the 5´ and 3´ termini of each DNA copy of each RNA molecule originally present in the sample – lends to our analyses high confidence about the fate of individual RNA molecules present in the starting material. Second, our data suggests a broad enrichment of fragments corresponding to histone mRNAs (Fig. 8a, red points). This observation is consistent with the work of others in defining the mode of translation initiation of histone mRNAs (at least H4 mRNA) to be independent of the m7G cap\(^{54}\). Interestingly, detailed biochemical characterization of H4 initiation by Martin et al. indicates that while the 5´cap is not required for 80S complex formation, it stimulates progression into the elongation phase of translation\(^{54}\). Thus, on our circRNA substrates containing histone mRNA fragments would be predicted to assemble initiation complexes that are unable to enter the elongation and termination phases of translation, perhaps indicating their strong degree of enrichment in our data.

**Complementary approaches to study IRES-mediated translation during hypoxia**

We have used two complementary approaches aimed at identifying physiological IRES usage in cells. Our ribosome profiling studies of cells exposed to a brief course of hypoxia defined which mRNA targets are significantly stimulated under a specific physiologic condition where translation is globally reduced\(^{60}\). Because of the short duration of hypoxia (4 hours), it is likely that the observed changes we observed occur at the level of translation rather than changes in RNA abundance. We identified specific upregulation of several key hypoxia-responsive mRNAs such as *EGLN1* and 2 (PHD1/2), *MTOR*, *HSPA1A* (Hsp70), and *VEGFA* (Fig. 9e). The *EGLN* genes, for example, encode a
set of prolyl hydroxylases that are the key sensors of molecular oxygen in metazoans that signal downstream to hypoxia-inducible factor-1 (HIF-1)\(^{56,57}\). Interestingly, the \textit{EGLN} genes are also involved in the adaptation of physiological traits needed for Tibetans to live at high altitude\(^{61}\). The mammalian target of rapamycin (mTOR) is also critical for mediating the effects of hypoxia on mammalian cells\(^{62}\). For example, hypoxia-dependent inhibition of oxidation phosphorylation leads to ATP depletion\(^{63-65}\), and causes mTOR-dependent changes in the phosphorylation of critical mTOR target 4E-BPs 1-3, eEF2, and RPS6K\(^{66}\), which serve to globally reduce protein synthesis by both reducing the availability of eIF4F (limiting new initiation) and reducing the activity of eEF2 (blocking translation elongation)\(^{49,67}\). The upregulation of MTOR expression by hypoxia in HEK293 likely represents cellular adaptation to low oxygen levels, as mTOR activity promotes continued growth\(^{68}\). Consistent with this observation, transformed cells lines (such as HEK293 cells) are notably more resistant to reduced oxygen levels than non-transformed cells\(^{66,67}\). It should be noted, however, that apoptosis of HEK293 cells occurs during prolonged (\(>4-6\) hours) of culture in arguably mild (1% \(O_2\)) hypoxia, indicating only modest hypoxic-resistance. In contrast, tumorigenic samples, particular those notably for growth in anoxic (0-0.1% \(O_2\)) microenvironments such as U87-MG glioblastoma cells, are essentially entirely resistant to the lack of oxygen.

\textit{Considerations involved in circular RNA expression}

In our high throughput screening design, we have included a strong translational pause site just before the luciferase stop codon (\textbf{Fig. 5a}). The purpose of this pause site is to increase the dwell time of translating ribosomes on the circRNA reporter and thus
increase the sensitivity of our profiling assay in enriching for IRES fragments. However, when express stall-containing circRNAs in cells, it may be prudent to consider the possibility that stalled ribosomes may recruit endonucleases resulting in circRNA cleavage\textsuperscript{69}. Further experiments which remove this translational pause site from the circRNA reporter, or performing IRES profiling in cells lacking \textit{PELO/HBSIL} may be useful next steps. Secondly, our cytosolic T7-dependent transcription strategy necessarily produces transcripts with 5´triphosphates eliciting cellular anti-viral responses\textsuperscript{21} and could limit reporter expression, and bias biological responses, in sensitive cell types. Finally, the expression of circular RNAs in cells could be broadly useful outside their utility in blocking cap-dependent translation, for example in the study of endogenous circRNAs\textsuperscript{70-74} which compete with pre-mRNA splicing\textsuperscript{73} and which may act as microRNA sponges regulating gene expression\textsuperscript{71}.

**Methods**

**T4 group I intron ribozyme sequences** Primers for amplifying the T4 group I intron (Thymidylate synthase) half ribozymes were obtained from Ford and Ares (1994)\textsuperscript{46}. T4 phage culture was obtained from ATCC (Cat. No. 11303-B39).

3´-intron

\[
\text{AATTCGGTTTCTACATAATGCCTAAACGACTATCCCTTTGGGAGTGGTTCAAGTGAACCTCGAA}
\text{ACGATAGACAACTTCTTTAAACAAGTTGGAGATATAGTCTGCTCTGCCATGGTGACATGCAGCT}
\text{GGATATAATTCCGGGGTAAGATTAACGACCTTATCTGAACATAATGCTACCCG}
\]

5´-intron

\[
\text{TCTTGGGT} \text{TAATGGGCGCTTGAATAGGGTACTTTATACCTTGTAATCTATCTAAACGCCCCTACCCGCTCCTAATTGGAT}\text{CTACCCG}
\]

5´-3´ splice junction

\[
\text{5´----TCTTGGGTCTACCCG---3´}
\]
5′intron (Δ18 splicing mutant; lacks 5′splice site)
GAGTATAAGGTGACTTATACTTGTAATCTATCTAAACGGGGAACCTCTCTAGTAGACAATCCC
GTGCTAAATTGTAGGACT

**In vitro transcription of circular RNAs** pEM plasmids were linearized with SspI and phenol-chloroform extracted. Aliquots were run on a 1% agarose gel to ensure complete digestion. 50-100 µl *in vitro* transcriptions were set up using Ambion T7 ultra IVT kit according to the manufacturer’s protocol. IVT reactions were carried out at 37°C for 1-2 hours. Transcripts for circularization were not capped or polyA tailed. Efficient circularization occurred co-transcriptionally and did not require additional incubations in high Mg<sup>2+</sup> solutions as described in Ford and Ares (1994) 46, and which in our hands led to higher rates of circRNA cleavage. IVT reactions were stopped by moving to ice and cleaned up by acid phenol/chloroform extraction followed by ethanol precipitation. Analysis of circular products may be performed by electrophoresis using a low percentage (<6%) TBE-Urea gel and formamide denaturing loading dye. RnaseR was used to destroy residual unreacted transcripts or cleaved circRNAs.

**In vitro translation (IVTL) of circular RNAs** Circular RNAs were used to program *in vitro* translation reactions in rabbit reticulocyte lysate (RRL). 1 µg of digested linear DNA was added to reactions having a final volume of 50 µl. IVTL reactions was allowed to proceed for 25 minutes at 37°C before adding cycloheximide (100 µg/ml final concentration) (Sigma) and snap freezing on liquid nitrogen.
Circular RNA expression in HEK293 cells

For expression of circRNAs in HEK293 cells, pEM-series plasmids were transiently transfected using Mirius-293 transfection reagent (Mirius) along with T7 RNA polymerase (pNTI-228). Cells were harvested 24-48 hours post-transfection and lysates were prepared by adding an equal volume of polysome buffer, vortexing for 10-20 seconds, and incubating on ice for 10 minutes. Lysates were clarified by centrifugation at 20,000xg for 10 minutes at 4°C in a refrigerated centrifuge.

Input library preparation for IRESeq

1-2 µg of total human brain RNA (Ambion) was used as starting material with the stranded TruSeq library prep kit (Illumina). The final PCR amplification step was skipped, and the splayed-end un-amplified library was instead used for custom primer PCR amplification using:

NI-EM-165 (5’-CAAGCAGAAGACGGCATACGA-3’)
with $T_{\text{annealing}} = 64^\circ$ and $T_{\text{extension}}=1:45$ seconds.

Gibson assembly of ribozyme halves and TruSeq library

Gibson assembly was performed using 30ng of custom TruSeq amplification library (amplified by 17 cycles) and cleaned-up using the DNA clean and concentrator column (Zymo) with 0.1 pmols of upstream and downstream reporter fragments.

The upstream reporter, containing the 3´-intron was amplified from pEM-230 using:

NI-EM-95 5´-CAGGTCGACTGTCGGGGCTGGCTAACA-3´
NI-EM-167 5´-GTTCGTCTTCTGCGTGCGCGTCCTAGTTCA-3´
The downstream reporter, containing the 5´-intron was amplified from pEM-230 using:

NI-EM-182: 5’-TGGTCGCCGTATCATTTAGCCACCATGGCTGCAGCTTGGGTTAA-3’
NI-EM-134: 5’-AGTCTACAATTTAGCAGGGA-3’

Assembly reactions were performed at 50°C for 1 hour according to the manufacturer’s protocol (NEB).

**Library amplification of Gibson Assembled Ribozyme-library** 5µl (50%) of each Gibson Assembly reaction was amplified by PCR for between 7 and 27 cycles using:

NI-EM-95 (5’-CAGGTCGACTGTCGGGGCTGGCTTAACTA-3’) and
NI-EM-134 (3’-AGGGCACGATTTAACATCCTGA-5’)

with $T_{\text{anneal}}=62$°C and $T_{\text{extension}}=2:00$ minutes. PCR reactions were cleaned up using a DNA clean and concentrator column (Zymo).

**Ribosome pelleting, RNA extraction, and reverse transcription of ribosome-associated circRNA** Ribosomes, and associated circRNAs, were prepared from clarified lysates (from HEK293 cells) or RRL IVTL reactions by pelleting over a 34% sucrose cushion using a tabletop ultracentrifuge and a TLA100.3 rotor. Samples were spun at 4°C for 1:00 hour at 100,000 rpm. Input or pelleted material was extracted 2X with acid phenol/chloroform and ethanol precipitated. Total RNA was quantified using the NanoDrop (thermo) or BioAnalyzer (Agilent). Reverse transcription reactions were assembled using 1µg of total RNA and 6µM random hexamers (NEB) using Superscript
II (Invitrogen) or ProtoScript (NEB). After reverse transcription, RNA was destroyed by added 1N NaOH and heating to 98˚C for 20 minutes.

**Circle-specific PCR using first strand cDNA** First strand cDNAs from ribosome pelleted material was amplified by circle-specific PCR prior to Illumina sequencing. Tanneal=56˚C with a 1 minute extension time for 25 cycles.

NI-NI-38 and NI-EM-264 (5´-AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CAC ACC TCC CTG TTC AAG GAC -3´)

**Circularization Plasmids**

*(a) General Reporter Structure:*

\[ pCMV-pT7-3i-dATG-NLuc-Xbp1CT-IdxAd-IRES\_Fragment-ATG- UnivAd-linker-FLAG-IgD-5i \]

*(b) IRES fragments*

*Cloned from human brain cDNA*

- pEM-231: HSPA1A (sense)
- pEM-232: HSPA1A (antisense)
- pEM-233: SNRNP_ICS (sense)
- pEM-234: SNRNP_ICS (antisense)
- pEM-236: HSPA5 (sense)
- pEM-237: HSPA5 (antisense)
- pEM-238: KCNK12 (antisense)
**Viral IRES sequences**

pEM-230: EMCV\(^{WT}\)

pEM-235: HCV (sense)

**pEM-230 Plasmid Features**

- **pCMV**: CMV promoter sequence 10:562
- **pT7**: T7 promoter sequence 622:638
- **3i**: 3'-intron of self-splicing T4 group I intron (Thymidylate Synthase) 644:821
- **dATG-NLuc**: NanoLuciferase ORF with AUG codon removed 828:1337
- **XBP1Ct**: C-terminal peptide stalling motif 1341:1421
- **IdxAd**: Illumina indexed adaptor sequence 1425:1488
- **EMCV_IRES-AUG**: with AUG codon at the 3' end 1489:2079
- **UnivAd-Linker**: 24 nucleotide (truncated) Illumina universal adaptor sequence in-frame with IRES_Fragment AUG 2080:2103
- **FLAG**: in-frame FLAG affinity tag 2104:2127
- **IgD**: in-frame IgG domain linker 2155:2328
- **5i**: 5'-intron of self-splicing T4 group I intron (Thymidylate Synthase) 2351:2451

**pEM-230 sequence**

```
TAGTTATATAGTATAATCATTACGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATACAATTACGGTA
AATGGCCCGCTTGGCTAGCCCGCAACAGACCCCGCCCGCTTGGCTTATTGACGTATGTCCTCATGTGACAGGGTGGAGTATTT
TAGGGACATTCACTGCAATGGGTCAGTTACGCTAGCTTGGCAGTACATCAAGTGTATCATATGCTATGCTAGCTTGGCAG
TGGATAGCGGTGATGCTCAGGCGGGATCGCTTACATCTCAGCAGGCTGGACATCATGCAAGGCTGGCTATATTCCGGGGTA
TTAACGACCTTATCTGACAGGTGATGATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTAC
GCCGAACATACIATTATAGTATAATCATTACGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATACAATTAC
```

155
Figure List

Figure 1. Cap-independent translation assay using a covalently circular RNA

Figure 2. Circular RNA IRES reporter for library screening
Figure 3. High throughput IRES profiling strategy

Figure 4. *in vitro* ribozyme-mediated cyclization

Figure 5. Nascent chain-mediated translational pausing in IRES reporter

Figure 6. Functional validation using the VEGFA IRES *in vitro* and in HEK293

Figure 7. Ribozyme-mediated IRESeq library cyclization and expression in HEK293 cells

Figure 8. IRESeq enrichment for specific library sequences

Figure 9. Translational control during hypoxia

Figure S1. IRESeq input library complexity

Figure S2. IRESeq spike-in control sequences

Figure S3. Translational changes during hypoxia and reoxygenation

**Figure Legends**

**Figure 1. Cap-independent translation assay using a covalently circular RNA reporter** | **a.** Cap-dependent initiation of translation involves the eIF4F cap binding complex, PABP, and 43S scanning complexes (left). IRES-mediated initiation bypasses the requirement for one or more initiation factors such as eIF4F leading to initiation independent of the 5´end of the mRNA (right). **b.** The translation of a covalently circular RNA reporter is translated only if it contains an IRES.

**Figure 2. Circular RNA IRES reporter for library screening** | **a.** Linear pre-spliced library structure which undergoes ribozyme-mediated intramolecular circularization. 5´and 3´intron (red) flank the sequence to be circularized in a permuted orientation (the
5′ ribozyme half is downstream while the 3′ribozyme half is upstream). After splicing, this results in the library-AUG-splice junction-ORF-STOP orientation of the reporter.

Figure 3. High throughput IRES profiling strategy | a. The 5′ and 3′ribozymes are first assembled with transcriptomic RNA fragments and a T7 promoter (pre-IRESeq library) (i). Along with T7 RNA polymerase, this library is then transfected into HEK293 cells (ii), where it undergoes T7-mediated cytoplasmic transcription and circularization (iii). Ribosomes, and any associated circular RNA, are then pelleted through a sucrose cushion (iv). Total RNA extracted from input and pelleted material is then used for reverse transcription using random hexamers. To ensure analysis of only circular RNA, to eliminate the contribute of RNA circles which may have been cleaved during pelleting, libraries are amplified by a circle-specific PCR strategy (v).

Figure 4. in vitro ribozyme-mediated cyclization | a. qRT-PCR-based strategy for quantification of circular RNA prepared co-transcriptionally from T7 in vitro transcription reactions and reverse transcribed by random hexamers. Inward (blue) or outward (red) -facing PCR primers are indicated. b. RT-PCR gel demonstrates ribozyme-mediated intramolecular circularization. Importantly, transcripts lacking the 5′splice site do not circularize. Oligo\textsuperscript{ctrl} reactions were performed using DNA templates which encode the predicted post-spliced sequence. c. qRT-PCR quantification of splicing efficiency using the wild type ribozyme or a splicing mutant that lacks that 5′-splice site. 1.6 and 3.0kb are two different sized transcripts used for circularization and indicate that larger insert sizes lead to modestly reduced circularization efficiency d. in vitro transcribed self-
splicing RNAs were run on a 6% TBE-Urea gel. The presence of the 5´ splice site correlates with the appearance of the removed 5´ and 3´ intron halves (d) as well as the appearance of an RNaseR-resistance population of RNA circles (e).

**Figure 5. Nascent chain-mediated translational pausing on circRNA reporter | a.**

Ribosome profiling of circRNA reporter used in rabbit reticulocyte *in vitro* translation. The XBP1 C-terminus (red) fused in-frame to the firefly luciferase coding sequence (blue) causes strong ribosome pausing at the stop codon.

**Figure 6. Functional validation using the VEGFA IRES *in vitro* and in HEK293 cells | a - b.**

circRNA reporters containing the VEGFA IRES or a non-IRES control fragment were translated in rabbit reticulocyte lysates or HEK293 cells. Ribosomes, and any associated circRNA, were pelleted through a sucrose cushion and analyzed by RT-PCR. NanoLuc primers amplify a fragment of the reporter independent of intramolecular circularization. Splice junction primers amplify a product only from productively circularized RNA template. c. Spliced products containing the IRES or control fragments were detected in HEK293 cells at similar levels. Analysis of the sequence of the spliced junction by TOPO cloning confirmed that 10/10 clones carried the predicted spliced junction sequence.

**Figure 7. Ribozyme-mediated IRESeq library cyclization | a.**

Self-splicing circularization library design strategy using splayed-end (un-amplified) Illumina TruSeq library and Gibson assembly to the constant linker regions. b. RNA library prepared by
Gibson assembly run on a 6% TBE-Urea gel showing efficiency self-splicing of library fragments. c. Expression of circular RNA libraries in HEK293 cells were analyzed 24 hours after transfection. Total (linear + circular) reporter RNA level (left panel) and protein level (center panel) was T7-dependent. circRNAs were detected in both input fractions as well as ribosome pellets (right panel).

**Figure 8. IRESeq enrichment for specific library sequences** | a. IRESeq libraries prepared from brain mRNA fragments and screened for IRES activity in HEK293 cells using spike-in control IRES sequences (See Fig. S2). Control sequences were efficiency recovered from the reaction (VEGFA, HSPA1A, and HSPA5) in addition to several unique library fragments, including those from histone genes which translated through a cap-independent (or cap-assisted) mechanism. b. Gene model for VEGFA translation using an IRES and its stimulation during hypoxia. IRESeq fragments recovered from the endogenous VEGFA5 5’UTR (red) were strongly enriched above those present in the rest of the mRNA in the input library (green). Matched ribosome profiling of HEK293 cells harvested after 4 hours in 1% O2 shows that VEGFA expression is stimulated by hypoxia and ribosomes cover the putative IRES region of the 5’UTR.

**Figure 9. Translational control during hypoxia** | a. HEK293 cells were grown in replicates under standard conditions in 20% (ambient) O2. 24 hours after transfection with the pre-IRESeq library, cells were either switched to 1% O2 or left in 20% O2 for 4 hours. At this time, cells were removed from the incubator and cycloheximide was added to the media while cells were spin down for lysis. b-d. gene models for translation during
normoxia (red) or hypoxia (blue). e. Volcano plot showing the fold change in ribosome protected fragments (RPFs; x-axis) and statistical significance (p-value; y-axis) following hypoxic treatment. Hypoxia-induced genes are indicated (blue dots) and significant thresholds are shown.

**Figure S1. IRESeq input library complexity | a-b.** The size distribution and library complexity of starting material for IRESeq was determined by RNA sequencing and monitored after each library manipulation. The number of unique molecules predicted based on sampling of sequencing reads was calculated using the PreSeq package. Using these quantitations, we estimate the starting material contained between $10^8$ and $10^9$ unique fragments and between $10^7$ and $10^8$ unique fragments after Gibson assembly, transcription, and circularization. c. Replicate scatter plots of circRNA expression in the supernatant of HEK293 cells and genomic regions present in the supernatant library (d). Paired-end sequencing defined the fragment size distribution of the final library to be screened by IRESeq as well as the size distribution of ribosome-associated circRNAs. Representative IRES sizes (for the MYC and VEGFA IRESes; in nucleotides) are indicated.

**Figure S2. IRESeq spike-in control sequences | a-b.** A test suite of known IRES fragments was assembled using each sequences antisense fragment as a negative control. This suite contained both known human and known viral IRESes ($EMCV^{\text{wt}}$, $EMCV^{\text{loss-of-function}}$, and HCV).
Figure S3. Translational changes during hypoxia and reoxygenation | a. MA plots of RPF fold changes after 4 hours of hypoxia (a) and 30 minutes after reoxygenation (b). Statistical significance is indicated as red dots using 1% false-discovery rate (FDR). Differential expression was calculated using the DESeq R package.
References


Fig. 1

(a) Cap-dependent initiation
(b) IRES-mediated initiation
Fig. 2

a

5' 3'-intron orf Library 5' intron UAA AUG 3'

↓

Splice Junction
Fig. 4

a

Circle primers

Total primers

b

RT-PCR

5′-splice site: + -

Total + -

Circular + -

Oligo<sup>Ctrl</sup>

![RT-PCR images]

RNA levels

Circularization efficiency

1.6kb 3.0kb Splicing mutant

Total Circular

1.2 1.0 0.8 0.6 0.4 0.2 0.0
Fig. 4 continued

**d**

*in vitro* transcribed RNA

<table>
<thead>
<tr>
<th>5′-splice site:</th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circular product</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear precursor</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Kb

<table>
<thead>
<tr>
<th>3′-intron</th>
</tr>
</thead>
<tbody>
<tr>
<td>5′-intron</td>
</tr>
</tbody>
</table>

**e**

RnaseR

<table>
<thead>
<tr>
<th>Time (hr):</th>
<th>1.5</th>
<th>2.0</th>
<th>1.5</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>5′-splice site:</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Kb

| 1.0 | 0.3 | 0.2 |

172
Fig. 5

a

circRNA reporter

orf  stall  IRES  ATG

Footprints

Nucleotides of transcript
**Fig. 6**

**a**

<table>
<thead>
<tr>
<th>RT-PCR</th>
<th>Rabbit reticulocyte Ribo-Pellet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>NanoLuc</strong></td>
</tr>
<tr>
<td></td>
<td>IRES</td>
</tr>
</tbody>
</table>

**b**

<table>
<thead>
<tr>
<th>HEK293T Ribo-Pellet</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NanoLuc</strong></td>
</tr>
<tr>
<td>IRES</td>
</tr>
</tbody>
</table>

**c**

10/10 TOPO-cloned splice-junctions had predicted sequence
Fig. 7

(a) Splayed-end Illumina TruSeq library

(b) Gibson RNA Library

IVT

+

Circular library
Precurser library
Cyclic 3'-intron
3'-intron
5'-intron

XXX = antisense
YYYY = sense
Fig. 7 continued

C

HEK293 cells

qPCR

Luciferase Activity

circRES

<table>
<thead>
<tr>
<th>Norm. RNA levels</th>
<th>NonTx</th>
<th>T7-</th>
<th>T7+</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>RLUs / 10^3</th>
<th>NonTx</th>
<th>T7-</th>
<th>T7+</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Norm. RNA levels</th>
<th>Input circRNA</th>
<th>Ribo-PUlset circRNA</th>
</tr>
</thead>
</table>

176
Fig. 8 continued

b

VEGFA

<table>
<thead>
<tr>
<th>IRESseq</th>
<th>Normoxia</th>
<th>Hypoxia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>pre-IRESseq library</th>
<th>Normoxia</th>
<th>Hypoxia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ribosome Profiling</th>
<th>Normoxia</th>
<th>Hypoxia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Translation

- IRES
- CUG
- AUG
- 5' 3'
Fig. 9 continued

Ribosome Profiling

HSPA1A (Hsp70)
EGLN1 (PHD1)
MTOR
VEGFA

Significant

Not Significant

Statistical Significance
[-log(p-value)]

Change in Hypoxia

[log2(Fold Change)]

Lower in Hypoxia
Higher in Hypoxia
Fig. S2

a

3'-intron orf IRES (sense) orf 5'-intron
5' 3'-intron orf IRES (antisense) orf 5'-intron
3'

b

IRES -AUG1-FlowCell1-AUG2-5'ss-3'ss-ORF-Stall-UAA

<table>
<thead>
<tr>
<th>Sense (+)</th>
<th>Antisense (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VEGFA^B-IRE S</td>
<td>VEGFA^A-IRE S</td>
</tr>
<tr>
<td>VEGFA^A-IRE S</td>
<td>VEGFA^A-IRE S</td>
</tr>
<tr>
<td>BIP</td>
<td>BIP</td>
</tr>
<tr>
<td>Hsp70</td>
<td>Hsp70</td>
</tr>
<tr>
<td>SNRPN ICS</td>
<td>SNRPN ICS</td>
</tr>
<tr>
<td>EMCV-wt</td>
<td>EMCV-wt</td>
</tr>
<tr>
<td>EMCV-lof</td>
<td>EMCV-lof</td>
</tr>
<tr>
<td>HCV</td>
<td>HCV</td>
</tr>
</tbody>
</table>

human - viral
Fig. S3

a

4 hrs Hypoxia

Log2FoldChange

Expression Level

b

4 hrs Hypoxia + 30 min Reoxygenation

Log2FoldChange

Expression Level
Curriculum Vitae

Eric W. Mills
e@mills12@jhmi.edu

Academic Education

2017  - expected -  M.D. Johns Hopkins University School of Medicine
2017  - expected -  Ph.D. Johns Hopkins University School of Medicine
               Dept. of Molecular Biology and Genetics

2009       B.S. Neuroscience, with High Honors
               University of Michigan, Ann Arbor, Michigan

Research Experience

2011-2015  Ph.D. Dissertation Research
               Dept. of Molecular Biology and Genetics, Johns Hopkins University School of Medicine (2013-2015) &
               Advisors: Nicholas Ingolia and Rachel Green
               Topic: Translational control mechanisms in hematopoietic cells

2007-2009  Undergraduate Thesis Research
               Dept. of Molecular, Cellular, and Developmental Biology,
               University of Michigan (Ann Arbor)
               Advisor: Haoxing Xu
               Topic: Intracellular transient-receptor potential (TRP) ion channel physiology and iron metabolism in neurodegeneration

Awards/Honors/Fellowships

2012-2014  American Heart Association predoctoral fellowship
               Title: Post-transcriptional control of gene expression in platelets

2010       Medical student research award (Johns Hopkins)
2008-09 Undergraduate Research Fellowship in the Dept. of Molecular, Cellular, and Developmental Biology (University of Michigan)
2008 Cullen trust for higher education award

Publications


Other publications
1. MedIQ Learning, LLC. Eric Mills, Invited co-author for the USMLE-Rx Question Bank 2013

Abstracts (†Chosen for oral presentation)


**Oral Seminars**

1. Biochemistry and Molecular Biology (BCMB) graduate program retreat, 2014 (Johns Hopkins)

2. Dept. Molecular Biology and Genetics Colloquium 2014 (Johns Hopkins)

3. Dept. Molecular Biology and Genetics Journal Club, May 2014 (Johns Hopkins) †Awarded journal club presentation distinction


10. Johns Hopkins University, Baltimore, Maryland. RNA Club Aug 2012