Control of human hemoglobin switching by LIN28B-mediated regulation of BCL11A translation

Anindita Basak, Mathias Munschauer, Caleb A. Lareau, Kara E. Montbleau, Jacob C. Ulirsch, Christina R. Hartigan, Monica Schenone, John Lian, Yaomei Wang, Yumin Huang, Xianfang Wu, Lee Gehlke, Charles M. Rice, Xiuli An, Helen A. Christou, Narla Mohandas, Steven A. Carr, Jane-Jane Chen, Stuart H. Orkin, Eric S. Lander and Vijay G. Sankaran.

Increased production of fetal hemoglobin (HbF) can ameliorate the severity of sickle cell disease and β-thalassemia. BCL11A represses the genes encoding HbF and regulates human hemoglobin switching through variation in its expression during development. However, the mechanisms underlying the developmental expression of BCL11A remain mysterious. Here we show that BCL11A is regulated at the level of messenger RNA (mRNA) translation during human hematopoietic development. Despite decreased BCL11A protein synthesis earlier in development, BCL11A mRNA continues to be associated with ribosomes. Through unbiased genomic and proteomic analyses, we demonstrate that the mRNA-binding protein LIN28B, which is developmentally expressed in a pattern reciprocal to that of BCL11A, directly interacts with ribosomes and BCL11A mRNA. Furthermore, we show that BCL11A mRNA translation is suppressed by LIN28B through direct interactions, independently of its role in regulating let-7 microRNAs, and that BCL11A is the major target of LIN28B-mediated HbF induction. Our results reveal a previously unappreciated mechanism underlying human hemoglobin switching that illuminates new therapeutic opportunities.

The developmental switch from fetal to adult hemoglobin in humans has been extensively studied and is of substantial interest for developing approaches to induce fetal hemoglobin (HbF) to treat sickle cell disease and β-thalassemia. Through functional and genetic follow-up of genome-wide association studies on HbF levels, BCL11A has been identified as a key regulator of both developmental hemoglobin switching and silencing of HbF in adults. BCL11A protein levels are developmentally regulated in humans such that, at the earlier developmental stages when HbF is highly expressed in erythroid cells, there is little or no BCL11A protein. In contrast, BCL11A protein is robustly expressed in adult erythroid cells that have low levels of HbF expression. Despite extensive studies, the basis of this developmental regulation of BCL11A protein expression—and thereby the upstream regulators of human hemoglobin switching—remains undefined.

In agreement with earlier studies, we found that BCL11A protein being more readily degraded in cells at the earlier developmental stages, as compared with adult erythroid cells. Since newborn erythroid cells showed some, albeit lower, expression of BCL11A protein, we could directly compare the rates of protein degradation by arresting protein synthesis with cycloheximide. No difference was observed in BCL11A protein degradation between newborn and adult erythroid cells (Fig. 1c,d). Thus, alteration in protein synthesis appeared to be a more likely mechanism to account for the variation in BCL11A protein levels. For direct measurement of protein synthesis rates, we metabolically labeled newly synthesized proteins with the methionine analog L-azidohomoalanine (L-AHA). Newborn and adult erythroid cells showed similar global L-AHA incorporation (Fig. 1e). By immunoprecipitation, we also noted similar levels of L-AHA-labeled GATA1 protein, a key erythroid transcription factor. In contrast, we observed a twofold reduction in newly synthesized BCL11A levels in newborn versus adult erythroid cells (Fig. 1f). Importantly, our immunoprecipitation procedure was effective at capturing nearly all detectable BCL11A and GATA1 protein from cell lysates (Extended Data Fig. 3a,b).
We reasoned that assessment of translation by examining the distribution of BCL11A mRNA across polysomes might allow us to identify mechanisms for the observed variation in protein synthesis rates.\(^1\)\(^2\)\(^6\). Surprisingly, we noted that BCL11A mRNA occupied actively translating ribosomes in polysome fractions with overlapping distributions in the newborn and adult erythroid cells (Fig. 1g,h and Extended Data Fig. 3c). This finding demonstrates that, while BCL11A had a lower rate of overall protein synthesis in newborn erythroid cells, its mRNA was bound by multiple ribosomes. To define more clearly whether there might be altered ribosome occupancy at specific regions of BCL11A mRNA, we conducted ribosome profiling.\(^1\)\(^2\)\(^6\)\(^7\). In agreement with our findings from polysome profiling, we found that the overall translational efficiency—a metric of the number of ribosomes bound across the entire mRNA normalized to total mRNA levels—was comparable in stage-matched newborn and adult pro-erythroblasts (Fig. 1f). However, specific regions of the BCL11A transcript showed reduced ribosome occupancy (Fig. 1f), white arrows), while other regions showed more consistent occupancy (Fig. 1f). This variation in the pattern of ribosome occupancy may suggest ineffective ribosome elongation as a cause of reduced protein synthesis in newborn erythroid cells.\(^1\)\(^2\)\(^6\)\(^7\). Reduced translational elongation can be mediated through interactions of an mRNA with RNA-binding proteins, as is seen with the fragile X mental retardation protein, FMRP.\(^6\)\(^1\)\(^6\)\(^1\)\(^9\)\(^2\)\(^9\). In an analogous manner, we reasoned that an RNA-binding protein expressed in erythroid cells at the earlier developmental stages might bind to the BCL11A mRNA and prevent its effective translation. In the case of FMRP, altered translation elongation is mediated through direct interactions with the ribosome.\(^2\)\(^6\)\(^1\)\(^9\)\(^2\)\(^9\). We therefore sought to identify ribosome-associated RNA-binding proteins in human erythroid cells that displayed appropriate developmental expression.

We conducted an unbiased survey of ribosome-interacting proteins in human erythroid cells through the use of the RNA antisense purification (RAP) method targeting the 18S ribosomal RNA coupled to quantitative mass spectrometry using stable isotope labeling by amino acids in cells (SILAC) (Fig. 2). \(^5\)\(^1\)\(^6\)\(^1\)\(^9\)\(^2\)\(^9\)\(^4\). Among the proteins most enriched in the 18S RAP experiment was the RNA-binding protein LIN28B, which showed enrichment similar to that of ribosomal proteins known to interact with the 18S rRNA (Fig. 2b,c and Supplementary Tables 1 and 2). This finding was validated through LIN28B co-immunoprecipitation (Fig. 2c). In cord blood-derived erythroid cells, a substantial fraction of LIN28B was localized to the cytoplasm and associated with translating ribosomes (Fig. 2d and Extended Data Fig. 4). Treatment of cell lysates with RNase A before polysome fractionation revealed some continued, albeit reduced, interactions of LIN28B with ribosomes, suggesting potential direct interactions beyond those that are RNA-dependent (Extended Data Fig. 5).

Importantly, LIN28B showed a gradient of expression across fetal liver, newborn and adult erythroid cells for both mRNA and protein levels (Fig. 2f,g, Extended Data Fig. 6a,b and Supplementary Table 2). LIN28B is an attractive upstream candidate regulator of BCL11A due to its documented role in binding numerous mRNAs and modulating their translation.\(^1\)\(^1\)\(^2\)\(^4\). Despite these studies, LIN28B has not previously been appreciated as a ribosome-associated protein. Motivated by the observed differences in LIN28B expression between fetal and adult hematopoietic cells, a recent study suggested that LIN28B may promote the reprogramming of hematopoietic progenitors into a fetal-like state and thereby elevate HbF levels, but no specific mechanisms for these activities were identified.\(^2\)\(^5\). In addition, LIN28B has been shown to be a key regulator at the early stages of developmental hematopoiesis.\(^2\)\(^6\)\(^2\)\(^9\).

We therefore tested whether the repression of BCL11A mRNA translation might be mediated by LIN28B. Initial experiments suggested that lentiviral-mediated LIN28B expression resulted in extremely high levels of LIN28B, far exceeding those seen physiologically, which was also the case in a previous study and which may have confounded this previous analysis.\(^6\)\(^1\)\(^9\)\(^2\)\(^9\). To obtain physiologically relevant LIN28B levels, we devised a strategy to sort cells based on expression of a green fluorescent protein (GFP) marker linked to the LIN28B complementary DNA transcript through an internal ribosomal entry site. By separating cells into GFP-high and GFP-low-expressing fractions, we were able to obtain differentiating hematopoietic progenitor cells that expressed LIN28B at levels comparable to those observed in normal physiology at the earlier stages of human erythroid development (Fig. 3a,b). While unregulated, high-level LIN28B overexpression showed a reduction in BCL11A mRNA and a concomitant upregulation of γ-globin (encoded by HBG1/2, Extended Data Fig. 7a–c), expression of LIN28B at physiologically relevant levels in adult erythroid cells resulted in an approximately twofold reduction in BCL11A protein levels with no appreciable change in its mRNA levels (Fig. 3b–d). Importantly, physiologically relevant expression of LIN28B was sufficient to robustly induce γ-globin (Fig. 3e) to an extent similar to higher-level expression and without perturbation of erythroid differentiation (Extended Data Fig. 3h).

While these results show that LIN28B can suppress BCL11A protein synthesis, the underlying mechanisms remained unclear, particularly given the pleiotropic functions of LIN28B. LIN28B can both negatively regulate the production of let-7 microRNAs and directly bind mRNAs with either a positive or negative impact on their translation (Fig. 3f).\(^2\)\(^6\)\(^2\)\(^9\)\(^4\)–\(^2\)\(^9\)\(^4\). Most studies on the role of LIN28B in hematopoiesis have focused on its role in let-7 microRNA biogenesis.\(^2\)\(^6\)\(^2\)\(^9\)\(^4\). Our observation of a LIN28B–ribosome interaction in erythroid cells suggests that let-7-independent roles might be important in this context. To decipher the mechanisms of LIN28B-mediated BCL11A regulation, we initially explored whether the
well-characterized let-7 microRNA pathway downstream of LIN28B might have a role in the regulation of human BCL11A expression and hemoglobin switching. Given that all let-7 family microRNAs harbor the same seed-targeting sequence (Fig. 3g), we introduced mimics of the let-7 microRNA let-7g into cord blood-derived erythroid cells (where let-7 is absent or weakly expressed while...
Fig. 2 | RNA-binding protein LIN28B associates with ribosomes in erythroid cells and is developmentally regulated. 

a, Schematic overview of RAP–mass spectrometry combined with SILAC mass spectrometry. 

b, Quantification of 18S and U1 interacting proteins. Scatter plot of log, transformed SILAC ratios from two biological replicates is shown. 

c, Immunoprecipitation (IP) in erythroid cells, using antibodies targeting LIN28B or control IgG. Immunoblot detection of ribosomal proteins RPL5, RPL11, RPS20, RPS19 and LIN28B, with GAPDH and β-actin as controls. –RNaseA and +RNaseA denote IP without or with RNaseA treatment for 30 min, respectively. Note gaps in the immunoblot between experimental conditions, which were created to ensure appropriate migration patterns for all proteins. Experiments were repeated two times independently. 

d, Representative confocal immunocytochemistry images of 4% paraformaldehyde-fixed newborn erythroid cells at differentiation day 7. The nuclear stain is DAPI (blue). LIN28B was detected with Alexa Fluor 488 (green), and tubulin (TUBA4A) was detected with Alexa Fluor 594 (red). Experiments were repeated five times independently. 

e, Representative immunoblots showing subcellular localization of LIN28B, Lamin B1 and GAPDH. The nuclear and cytoplasmic fractions are labeled. Experiments were repeated three times independently. 

f, LIN28B mRNA distribution (normalized to GAPDH expression) in fetal, newborn and adult erythroid cells at differentiation day 7 from HSPCs (n = 3 per time point; three biologically independent experiments). Mean is plotted and error bars show s.d. 

g, Representative immunoblots showing LIN28B expression in fetal, newborn and adult at differentiation day 7 from HSPCs (three independent experiments). Arrows indicate two LIN28B isoforms and asterisk indicates a non-specific band. β-Tubulin is used as a loading control. Blots have been cropped, and the corresponding full blots are available in the Source Data files. m/z, mass divided by charge number.
Fig. 3 | LIN28B alters BCL11A mRNA translation independently of the canonical let-7 microRNA pathway. a, LIN28B expression in adult cells separated into GFP-high and GRP-low populations \((n=3;\) three independent experiments). PE, phycoerythrin. b, LIN28B expression (normalized to GAPDH) in GFP-high and GAPG-low populations compared to physiological expression of LIN28B in fetal, newborn and adult cells \((n=3\) per time point; three biologically independent experiments). Mean is plotted, and error bars show s.d. Two-sided Student’s t-test was used. NS, not significant; \(P=0.052.\) c, Protein levels of LIN28B and BCL11A following low-level LIN28B expression \((n=3;\) three biologically independent experiments). Mean is plotted, and error bars show s.d. Two-sided Student’s t-test was used. ***\(P<0.001.\) d, Relative BCL11A mRNA levels \((n=3;\) three biologically independent experiments). Mean is plotted, and error bars show s.d. Two-sided Student’s t-test was used. ***\(P<0.001.\) e, Relative \(\gamma\)-globin expression as a percentage of total globins \((\gamma\text{- and } \beta\text{-globins})\) following low-level LIN28B expression on differentiation day 15 \((n=3;\) three biologically independent experiments). Mean is plotted, and error bars show s.d. Two-sided Student’s t-test was used. **\(P<0.01.\) f, Activities of LIN28B are mediated through either let-7 miRNAs, leading to repression of mRNA targets \((\mathcal{L})\), or direct mRNA binding, leading to activated \((\mathcal{D})\) or repressed \((\mathcal{R})\) translation \((\mathcal{T})\). g, Comparison of seed sequences for human let-7 microRNAs. h, HMGA2 mRNA levels \((n=3;\) three biologically independent experiments). Mean is plotted, and error bars show s.d. Two-sided Student’s t-test was used. ***\(P<0.001.\) g, Comparison of seed sequences for human let-7 microRNAs. h, HMGA2 mRNA levels \((n=3;\) three biologically independent experiments). Mean is plotted, and error bars show s.d. Two-sided Student’s t-test was used. ***\(P<0.001.\) i, Representative immunoblots showing BCL11A expression in newborn cells expressing GAPDH-targeting, scrambled or hsa-let-7g mimics on erythroid differentiation day 7 \((n=3;\) two independent experiments). j, The \(\gamma\)-globin levels following scrambled or hsa-let-7g mimic expression in newborn cells on day 15 of differentiation \((n=3;\) three independent experiments). Mean is plotted, and error bars show s.d. Two-sided Student’s t-test was used. NS, not significant; \(P=0.0597.\) k, HMGA2 mRNA expression \((n=3;\) three independent experiments). Mean is plotted, and error bars show s.d. Two-sided Student’s t-test was used. NS, not significant; \(P=0.052.\) l, HMGA2 mRNA expression following low-level LIN28B expression \((n=3;\) three independent experiments). Mean is plotted, and error bars show s.d. Two-sided Student’s t-test was used. ***\(P<0.001.\)
LIN28B is robustly expressed) and observed appropriate suppression of a known let-7 mRNA target, HMG12A (Fig. 3h). However, there was no change in BCL11A protein or γ-globin mRNA levels despite increased let-7 activity (Fig. 3i, j). To bolster these findings, we introduced LIN28B into adult hematopoietic progenitors and differentiated the cells toward the erythroid lineage. Concomitant introduction of let-7 mimics successfully suppressed the LIN28B-mediated upregulation of HMG12A mRNA, but failed to suppress the LIN28B-mediated induction of γ-globin (Fig. 3k, l). Together, these experiments show that LIN28B-mediated suppression of BCL11A protein expression and concomitant HbF induction are independent of let-7 microRNA activity.

An alternative possibility is that LIN28B directly influences BCL11A mRNA translation through direct interactions, as previously described for other mRNA targets. We therefore immunoprecipitated LIN28B from primary newborn erythroid cells after formaldehyde crosslinking and fractionation of the cells. We observed an enrichment of BCL11A mRNA relative to input, which was also seen for HMG1A, a known direct binding target of LIN28B, but not for several non-target mRNAs (Extended Data Fig. 8). However, formaldehyde crosslinking cannot discriminate direct interactions from indirect RNA–protein binding events. To overcome this limitation, we applied ultraviolet crosslinking and immunoprecipitation (CLIP) together with massively parallel sequencing to map endogenous LIN28B binding sites in newborn erythroid cells. We generated size-matched input libraries to calculate the enrichment of each peak over background. Using this approach, we identified 20,001 distinct binding sites of LIN28B that covered a total of 5,640 mRNAs in these cells at a 1% irreproducible discovery rate (Fig. 4a and Extended Data Fig. 9). LIN28B-bound mRNAs showed a range of translational changes when comparing ribosome profiles from newborn and adult erythroid cells (Extended Data Fig. 9f), consistent with previous studies demonstrating disparate modes of regulation among LIN28B-bound mRNAs. Notably, we identified a 139-base pair (bp) LIN28B binding site in the 3′ coding region of the BCL11A mRNA (Fig. 4b). This peak in BCL11A mRNA was among the top 1% of all observed binding peaks in newborn erythroid cells (Fig. 4a and Extended Data Fig. 7). This region contained three instances of a previously reported LIN28 binding motif (GGAG) and overlapped with the coding sequence that is unique to the XL isoform of BCL11A, which is critical for HbF silencing (Fig. 4c).

To assess the functional importance of the LIN28B binding site on BCL11A, we disrupted all three GGAG motifs in the LIN28B binding site of BCL11A cDNA by synonymous base changes (G>A),
which resulted in an increase in BCL11A protein synthesis in the presence of LIN28B expression (Fig. 4d). Importantly, further elevation of LIN28B expression further suppressed protein production from the wild-type (WT) but not the triple GGAG motif-mutated form of BCL11A (Fig. 4c). We have previously shown that the difference between mouse and human hemoglobin switching is mediated through alterations in BCL11A protein expression in definitive erythroid cells (Fig. 4f). Crucially, the first GGAG motif (termed site 1 here) displays a synonymous change in the orthologous position in mouse and human BCL11A mRNA (Fig. 4c, underlined). Strikingly, mutation of this single position (G>A) was sufficient to enable increased protein production of BCL11A in the presence of LIN28B (Fig. 4g), suggesting that this distinct region of BCL11A mRNA underlies the developmental differences between mice and humans in BCL11A protein expression and hemoglobin switching.

While these results emphasize a major role for LIN28B in the regulation of BCL11A translation and human hemoglobin switching, we sought to assess the degree to which the suppression of BCL11A mRNA translation results in the observed Hbf induction by LIN28B (Fig. 4h). When we simultaneously expressed LIN28B and BCL11A in primary adult hematopoietic progenitors undergoing erythroid differentiation, γ-globin levels were reduced to nearly the extent seen at baseline conditions in these adult erythroid cells (Fig. 4h) and Extended Data Fig. 10). In addition, we examined the levels of various let-7 microRNAs and observed no significant change after increasing the expression of BCL11A in the setting of LIN28B transduction, reinforcing the independence of the observed translational regulation from the well-characterized functions of LIN28B in regulation of let-7 family microRNA biogenesis (Supplementary Fig. 1).

Together, our results show that LIN28B is expressed in a developmental, stage-specific manner during the ontogeny of human erythropoiesis, such that at earlier developmental stages LIN28B is able to directly bind BCL11A mRNA and prevent its effective translation. The absence of LIN28B expression in adult erythroid cells allows effective BCL11A protein synthesis, and γ-globin expression is thereby suppressed. The observed graded translational inhibition of BCL11A mRNA by LIN28B thus promotes the gradual shift from fetal to adult hemoglobin expression during development17. Our findings provide important insights into the upstream regulators of hemoglobin switching, and suggest a simplified model for the regulation of this process. In addition, we show that alterations of the LIN28B regulatory axis by modulation of an individual evolutionarily divergent interaction site are sufficient to confer species-divergent BCL11A protein expression and hemoglobin switching18. Moreover, our study illuminates a key role for translational regulation of BCL11A in human development, beyond previously well-studied roles dependent on the let-7 microRNA16–29. This finding is one of the first examples of a physiologic function of LIN28B that are regulated in a similar manner and that may impact other development processes. This newly uncovered translational regulatory mechanism or upstream regulators of LIN28B may represent unique targetable opportunities for induction of HBF in the treatment of sickle cell disease and β-thalassemia.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-019-0568-7.

Received: 1 November 2018; Accepted: 11 December 2019; Published online: 20 January 2020

References
1. Sankaran, V. G. & Orkin, S. H. The switch from fetal to adult hemoglobin. Cold Spring Harb. Perspect. Med. 3, a011643 (2013).
2. Sankaran, V. G. et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science 322, 1839–1842 (2008).
3. Sankaran, V. G. et al. Developmental and species-divergent globin switching are driven by BCL11A. Nature 460, 1093–1097 (2009).
4. Rasak, A. et al. BCL11A deletions result in fetal hemoglobin persistence and neurodevelopmental alterations. J. Clin. Invest. 125, 2363–2368 (2015).
5. Liu, N. et al. Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. Cell 173, 430–442 e17 (2018).
6. Xu, J. et al. Transcriptional silencing of γ-globin by BCL11A involves long-range interactions and cooperation with SOX6. Genes Dev. 24, 783–798 (2010).
7. Martyn, G. E. et al. Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. Nat. Genet. 50, 498–503 (2018).
8. Sankaran, V. G. & Weiss, M. J. Anemia: progress in molecular mechanisms and therapies. Nat. Med. 21, 231–239 (2015).
9. Menzel, S. et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat. Genet. 39, 1197–1199 (2007).
10. Uda, M. et al. Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proc. Natl Acad. Sci. USA 105, 1620–1625 (2008).
11. Yen, H. C., Xu, Q., Chou, D. M., Zhao, Z. & Elledge, S. J. Global protein stability profiling in mammalian cells. Science 322, 918–923 (2008).
12. Ludwig, L. J. et al. Altered translation of GATA1 in Diamond-Blackfan anemia. Nat. Med. 20, 748–753 (2014).
13. Dieterich, D. G., Link, A. J., Graumann, J., Tirrell, D. A. & Schuman, E. M. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). Proc. Natl Acad. Sci. USA 103, 9482–9487 (2006).
14. Sankaran, V. G. et al. Exome sequencing identifies GATA1 mutations resulting in Diamond–Blackfan anemia. J. Clin. Invest. 122, 2439–2443 (2012).
15. Khajuria, R. K. et al. Ribosome levels selectively regulate translation and lineage commitment in human hematopoiesis. Cell 173, 90–103 e19 (2018).
16. Darnell, J. C. et al. FMRP stalls ribosomal translation on mRNAs linked to synaptic function and autism. Cell 146, 247–261 (2011).
17. Ingolia, N. T. Ribosome footprint profiling of translation throughout the genome. Cell 165, 22–33 (2016).
18. Das Sharma, S. et al. Widespread alterations in translation elongation in the brain of juvenile Fmr1 knockout mice. Cell Rep. 26, 3313–3322 e5 (2019).
19. Richter, J. D. & Coller, J. Pausing on polyribosomes: make way for elongation in translational control. Cell 163, 292–300 (2015).
20. Chen, E., Sharma, M. R., Shi, X., Agrawal, R. K. & Joseph, S. Fragile X mental retardation protein regulates protein translation by binding directly to the ribosome. Mol. Cell 54, 407–417 (2014).
21. McHugh, C. A. et al. The Xist IncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature 521, 232–236 (2015).
22. Rappold, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. 2, 1896–1906 (2007).
23. Munschauer, M. et al. The NORAD IncRNA assembles a topoisomerase complex critical for genome stability. Nature 561, 132–136 (2018).
24. Graf, R. et al. Identification of LIN28B-bound microRNAs reveals features of target recognition and regulation. RNA Biol. 10, 1146–1159 (2013).
25. Lee, Y. T. et al. LIN28B-mediated expression of fetal hemoglobin and production of fetal-like erythrocytes from adult human erythroblasts ex vivo. Blood 122, 1034–1041 (2013).
26. Yuan, J., Nguyen, C. K., Liu, X., Kanellopoulou, C. & Muljo, S. A. Lin28b reprograms adult bone marrow hematopoietic progenitors to mediate fetal-like lymphopoiesis. Science 335, 1195–1200 (2012).
27. Rowe, R. G. et al. Developmental regulation of myeloid erythroid progenitor function by the Lin28b–let–7–Hmgaa2 axis. J. Exp. Med. 213, 1497–1512 (2016).
28. Copley, M. R. et al. The Lin28b–let–7–Hmgaa2 axis determines the higher self-renewal potential of fetal haematopoietic stem cells. Nat. Cell Biol. 15, 916–925 (2013).
29. Bronnertsky, B. U., Turt, D. & McCune, J. M. Lin28b regulates fetal regulatory T cell differentiation through modulation of TGF-beta signaling. J. Immunol. 197, 4344–4350 (2016).
30. Newman, M. A., Thomson, J. M. & Hammond, S. M. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* **14**, 1539–1549 (2008).
31. Viswanathan, S. R., Daley, G. Q. & Gregory, R. I. Selective blockade of microRNA processing by Lin28. *Science* **320**, 97–100 (2008).
32. Wilbert, M. L. et al. LIN28 binds messenger RNAs at GGAGA motifs and regulates splicing factor abundance. *Mol. Cell* **48**, 195–206 (2012).
33. Hafner, M. et al. Identification of mRNAs bound and regulated by human LIN28 proteins and molecular requirements for RNA recognition. *RNA* **19**, 613–626 (2013).
34. Van Nostrand, E. L. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016).
35. Hafner, M. et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
36. Cho, J. et al. LIN28A is a suppressor of ER-associated translation in embryonic stem cells. *Cell* **151**, 765–777 (2012).
37. Nam, Y., Chen, C., Gregory, R. I., Chou, J. J. & Sliz, P. Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell* **147**, 1080–1091 (2011).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020
Methods

Isolation of CD34+ hematopoietic stem and progenitor cells (HSPCs). CD34+ HSPCs from various stages of human development were isolated by positive magnetic selection using the EasySep Human CD34 Positive Selection Kit II (Stem Cell Technologies) after mononuclear cell isolation on a Ficol–Paque density gradient (Stem Cell Technologies). The purity of isolated cells was assessed by flow cytometry with a physiologic expression conjugated anti-human CD34 antibody (Supplementary Table 3). Freshly isolated CD34+ HSPCs were either immediately cultured or cryopreserved for later use.

In vitro erythroid differentiation from primary hematopoietic cells. CD34+ HSPCs were cultured in a three-stage erythroid differentiation culture system previously described38. Briefly, cells were cultured in Iscove’s modified Dulbecco’s medium (IMDM, Life Technologies) containing 2% human AB plasma, 3% human AB serum, 1% penicillin/streptomycin, 200 μg/ml of human holo-transferrin, 3 μM of heparin and 15 μg/ml of insulin (base medium). In phase I (7–24 d), the cultures were supplemented with erythropoietin (EPO) (3 U/ml), stem cell factor (SCF, 10 ng/ml) and interleukin 3 (IL-3, 1 ng/ml), and in phase II (7–12 d) they were supplemented with EPO and SCF alone. In phase III (12 d onwards), primary cell cultures contained 1 mg/ml of human holo-transferrin supplemented with EPO. Cell counts were measured either using a Beckman Coulter automated cell counter or manually with a hemocytometer. Cells were maintained at 10^5–10^6/ml in phases I and II, and at 10^6/ml in phase III. Cells were split into fresh culture medium every 3 d.

Flow cytometry. Cells were washed in PBS and stained with anti-human CD71 and CD235a antibodies (Supplementary Table 3). Propidium iodide or DAPI was used for live/dead cell discrimination. Flow cytometric analyses were conducted on Becton Dickinson LSRII, LS Fortessa or Accuri C6 instruments and all data were analyzed using FlowJo software (v.10.3).

RNA-sequencing. CD34+ HSPCs from adult and newborn were differentiated until 7 d (primarily pro-erythroblasts and basophilic erythroblasts) and RNA was isolated. CDNA libraries were generated from 2 μg of RNA with TrueSeq RNA Sample Prep Kit v.2 (Illumina). Sequencing of 101 nucleotide paired-ends reads was carried out on an Illumina HiSeq 2000 sequencing system. Adapters were removed with trimmomatic using the following options: ‘PE -phred33 ILLUMINACLIP:TruSeq-PE-2.fa:2:30:10 HEADCROP:5 LEADING:10 TRAILING:10 SLIDINGWINDOW:4:20 MINLEN:36.’ Transcript abundance per million reads for each sample was estimated using kallisto v.0.4.4 using

RNA isolation and quantitative PCR with reverse transcription (RT–qPCR). RNA was isolated using the RNeasy Mini Kit (Qiagen) with on-column DNAse (Qiagen) digestion, according to the manufacturer’s instructions. RT–qPCR was carried out using a 96-well plate on a CFX96 Real Time System (BioRad) with iQ SYBR Green Supermix (BioRad)4. Gene-specific primers used for RT–qPCR are listed in Supplementary Table 4.

Immunoblot analysis. Cells were lysed in RIPA buffer (no. sc-24948, Santa Cruz Biotechnology) and protein levels were quantified with the DC Protein Assay on a CFX96 Real Time System (BioRad). Protein levels were quantified with the DC Protein Assay (BioRad). Cells were incubated with 100 μg/ml of cycloheximide (Sigma Aldrich) for 10 min at 37°C, washed twice with ice-cold PBS containing 100 μg/ml of cycloheximide and lysed in polysome extraction buffer (10 mM Tris-HCl, pH 7.4, 1% Triton X-100, 5 mM MgCl₂, 100 mM KCl, 2 mM DTT, 100 μg/ml of cycloheximide, 500 μM of RNAsin Plus (Promega) and complete EDTA-free Protease Inhibitor (Roche)), by shearing the lysate gently through a 26-gauge needle. The lysate was obtained after centrifugation for 10 min at 1,300 g. 4°C. Polymers were layered on a 10–50% linear sucrose gradient containing 20 mM HEPEs-KOH (pH 7.4), 5 mM MgCl₂, 100 mM KCl, 2 mM DTT, 100 μg/ml of cycloheximide and SUPERase In RNase Inhibitor (Ambion), and centrifuged at 36,000 rpm, for 2 h in a SW41 rotor using an XE-90 ultracentrifuge (Beckman Coulter). Proteins were fractionated in a BioComp Gradient Station fractionator, with absorbance at 254 nm used to visualize the gradients.

RNA was extracted on the collected fragments with TRIzol LS reagent (Ambion). Primers BCL11A, GATA1 and ACRB were used for the calculation of mRNA abundance in monosomic and polysome gradient fractions, as previously described19, and are represented as a percentage of the entire sucrose gradient. This representation was used for accuracy, since individual fractions did not have consistent volumes. When showing the overall pattern of the sucrose gradient fractions, the x axis is the total size (nm) of the sucrose fraction. Given the inability to collect identical fraction sizes for different samples, mRNA levels have been quantified to normalize the x axis as a percentage of total sucrose gradient.

In vitro protein labeling and detection. Metabolic labeling of proteins was carried out using Click-IT chemistry18. CD34+ HSPCs derived from adults and newborns were cultured in parallel until 7 d of erythroid differentiation, washed with warm PBS, and incubated in methionine-free RPMI medium to which 1% FBS and 2 mM l-glutamine for 1 h at 37°C, to deplete both cell types of methionine reserves. Methionine-starved cells were then labeled with Click-IT l-AHA (Life Technologies) at a concentration of 50 μM for 6 h at 37°C. Harvested cells were washed twice in PBS, warmed to 37°C and protein extraction was carried out in RIPA lysis buffer supplemented with phenylmethyl sulfonyl fluoride, sodium orthovanadate and protease inhibitor cocktail solution (Santa Cruz Biotechnology). Protein lysate (50 μg) was used for the Click reaction with tetrathyrrylrhodamine alkyl (TAMRA, Life Technologies), in a total volume of 200 μl with the Click-IT Protein Reaction Buffer Kit (Thermo Fisher Scientific). TAMRA signal was detected on total proteins separated by SDS gel electrophoresis with an Alphal imager (Amersham Biosciences).

Intracellular localization of BCL11A (Supplementary Table 3). Intracellular localization of cytokine receptors was analyzed by immunofluorescence using antibodies against CD71 and CD235a (Life Technologies) at a concentration of 50 μg/ml for 30 min.

Immunoprecipitation coupled to Click-IT labeling. Cell lysate obtained from Click-IT labeling was immunoprecipitated using a mouse monoclonal antibody against BCL11A (1A4S, no. ab19487, AbCam) and a goat anti-mouse (no. sc-1234, Santa Cruz Biotechnology) bound to Dynabeads Protein G (no. 10003D, Life Technologies) for 3 h at 4°C with end-to-end rotation. The Dynabeads–antibody–antigen complex was washed three times with RIPA buffer, resuspended in 60 μl of RIPA buffer and detected using the Click-IT Protein Reaction Buffer Kit (no. C10276, Invitrogen). The immunoprecipitated resin was washed twice with RIPA buffer, 100 μl of RIPA buffer, and bound proteins were eluted in 40 μl of Laemmli sample buffer (no. 161-0747, BioRad) and heated at 70°C for 10 min. Proteins were separated by SDS gel electrophoresis, followed by TAMRA detection as described above or immunoblot analysis using antibodies against BCL11A and GATA1. The immunoprecipitate fractions and flow-throughs were probed under identical experimental conditions (simultaneous SDS gel electrophoresis, same primary and secondary antibody dilutions used).
profiles. Cells were incubated with 100 μg/ml of cycloheximide (Sigma Aldrich) for 5 min at 37 °C, washed twice with ice-cold PBS containing 100 μg/ml of cycloheximide and lysed in 10 mM Tris-HCl (pH 7.4), 5 mM MgCl2, 100 mM KCl, 1% Triton X-100, 100 μg/ml of cycloheximide, 500 mM EDTA (Promega) and 1× Complete Protease Inhibitor, EDTA-free (Roche). Lysates for ribosome profiling were prepared, and ribosome-protected RNA fragments (RPFs) were isolated as described previously with the following modifications: ribosomal RNA was depleted using the RiboMinus Eukaryote Kit v2 (Thermo Fisher Scientific) following the manufacturer’s instructions. RNA-depleted RPFs were purified with the Zymo RNA Clean & Concentrator-5 kit following the manufacturer’s instructions for small RNA samples. PAGE purification of RPFs was performed as described previously, and resulting RPFs were dephosphorylated in FastAP mix (19.5 μl of H2O, 2.5 μl of 10X FastAP buffer (Thermo Fisher Scientific), 2.5 μl of FastAP enzyme (1 U/μl, Thermo Fisher Scientific), 0.5 μl of Murine RNase inhibitor (5 units/mL), 0.5 μl of Murine RNase (5 units/mL) and 0.5 μl of 10X PKN buffer (New England Biolabs), 0.5 μl of Murine RNase Inhibitor, 6.5 μl of T4 PNK (10 μl μl−1; New England Biolabs), 1 μl of TURBO DNase) with 75 μl added to each 25-μl sample and incubated for 20 min at 37 °C. RPFs were purified with the Zymo RNA Clean & Concentrator-5 kit following the manufacturer’s instructions for small RNA samples. RNA was eluted in 7 μl of H2O, then 1.5 μl of DMSO and 10 μl of 3′ adapter (Phos/AGATCGGAAGAGCACACGTCTG/ddC) were added before denaturation at 65 °C for 2 min and transfer of samples to ice. Subsequently, 11 μl of ligated mix was added (2 μl of 10X T4 RPF ligation buffer (New England Biolabs), 3 μl of 5X T4 DNA ligase buffer, 0.5 μl of T4 DNA ligase, 7 μl of 10X T4 DNA ligase High Concentration (New England Biolabs)) using low-retention pipette tips and incubated for 1 h at 25 °C with agitation. Ligation reactions were purified to remove free 3′ adapter using Silane bead purification. For each reaction, 15 μl of Silane beads (Thermo Fisher Scientific) were washed twice in 1 ml of RLT buffer (Quagen), beads were resuspended in 60 μl of RUT buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% SDS) and incubated at 65 °C for 2 min. Next, 0.7 volumes of 100% ethanol were added and incubated for 10 min at room temperature. Supernatant was removed and beads were washed twice with 70% ethanol before elution of air-dried beads in 10 μl of H2O. Next, 10 μl of reverse transcription primer (5′-biotin-GACGTGTGCTCCTGGA) was added and samples were denatured at 72 °C for 3 min before transfer to ice. To each reaction 7 μl of reverse transcription mix (0.25 μl of H2, 2 μl of 5× Smartscripse first strand buffer (Takara), 0.25 μl of RNase inhibitor, 2 μl of 50% PEG 8000 and 1.2 μl of T4 RNA Ligase 1 High Concentration (New England Biolabs)) were added. Samples were purified with Silane beads and subjected to PCR amplification using 2× NEBNext Q5 Hot Start HiFi PCR Master Mix (New England Biolabs) as described previously. Amplicons migrating at a size of ~157 bp were subjected to two consecutive rounds of gel purification. For preparation of mRNA-sequencing libraries, total RNA was extracted with the Direct-zol RNA MiniPrep Plus Kit (Zymo Research) according to the manufacturer’s instructions. Total RNA was poly-A selected using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs) according to the manufacturer’s instructions. mRNA was eluted in 27 μl of H2O, 3 μl of 10X FastAP buffer was added and RNA was heat-fragmented for 3 min at 91 °C. Fragmented RNA was cooled on ice and flash frozen with liquid nitrogen. All libraries were sequenced on a HiSeq 2500 system (Illumina). After recovery, cells were harvested on 7 d of differentiation for immunoblotting with the primary antibodies to vSVG, δκφ, δκφ and the laminogenomic vector of interest. The medium was changed to base medium on the day following transfection, and the viral supernatant was collected approximately 48 h post-transfection, filtered with a 0.45-μm filter and used for infection of CD34⁺ cells. Where the virus was required to be concentrated, 293 T cells were seeded in 10-cm dishes and the viral supernatant was filtered and centrifuged at 24,000 rpm for 2 h at 4°C. Between 200,000 and 300,000 CD34⁺ cells were infected in six-well plates with 8 μg/ml of polybrene (Millipore), spun at 2,000 rpm for 1.5 h at room temperature and incubated in the viral supernatant at 37°C. Virus was washed off 1 d after infection, and infected cells were selected for by GFP or RFP expression driven by IRES-GFP or IRES-RFP in the HMD vector in the particular construct. GFP⁺, RFP⁺ and GFP⁺/RFP⁺ cells were sorted by fluorescence-activated cell sorting (FACS) in a sterile manner and cultured for further analysis.

Transfections in 293T cells. The 293T cells were maintained in DMEM with 10% FBS and 1% penicillin/streptomycin. Lentiviral infections were carried out as described above. Transfection of constructs was carried out using 1 μg of the vector of interest with OPTI-MEM (Gibco) and Fugene 6 (Promega).

Introduction of mimics. mirIDIAN microRNA mimics for human let7g (no. CP-001000-05-0005, Dharmacon) were used, along with positive (no. CP-001000-02-05, Dharmacon) and negative (no. CP-002000-01-05, Dharmacon) controls. The Amaxa human CD34⁺ Cell Nucleofector Kit (no. VPA-1003, Lonza) was used to electroporate 50 nM of each mimic into CD34⁺ cells derived from newborns on 5 d of erythroid differentiation, according to the manufacturer’s instructions. After recovery, cells were harvested on 7 d of differentiation for immunoblotting and RT-qPCR, as described above. Small interfering RNA mimics targeting GAPDH, human let7g and a scrambled negative control are shown.

Immunocytochemistry analysis. Differentiating HSPCs derived from newborns on 7 d of differentiation were cultured with 4% paraformaldehyde and fixed with Triton X-100. Let7g antibody (no. ab150077, AbCam, 1/1,000 dilution) and Alexa Fluor 488-conjugated goat anti-rabbit IgG secondary antibody (no. ab24170, AbCam, 1/1000 dilution) were incubated with cells for 1 h at 4°C. After washing the cells, secondary IgG Alexa Fluor 488 goat anti-mouse secondary antibody (no. 24170, AbCam, 1/1000 dilution) was added, and slides were observed under an Olympus Flouview FV1000 confocal microscope.

Subcellular localization studies. Newborn-derived HSPCs were cultured until 7 d of differentiation. Nuclear and cytoplasmic fractionation was performed with the PARIS kit (no. AM1921, Ambion) according to the manufacturer’s instructions. Immunoblots were performed with antibodies LIN28B, GAPDH (described above) and Lamin B1 (B10, no. sc-374015, Santa Cruz, 1/1,000 dilution).

Statistical analysis. All statistical significances between control and test groups were calculated with two-tailed Student’s t-test. Biological replicates are mentioned.
wherever applicable. Error bars show s.d. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001; NS, not significant throughout the text. All statistical analyses were carried out in GraphPad Prism 7.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**
The massively parallel sequencing data associated with this manuscript are available in the Gene Expression Omnibus database under accession code GSE118359. The original mass spectra and the protein sequence database used for searches have been deposited in the public proteomics repository MassIVE (http://massive.ucsd.edu) and are accessible at accession MSV000084443. Source data for Figs. 1–4 and Extended Data Figs. 1, 3–5 and 10 are available online.

**Code availability**
Custom computer code for reproduction of sequencing-based analyses is available at https://github.com/sankaranlab/translation-regulation-bcl11a.

**References**
38. Giani, F. C. et al. Targeted application of human genetic variation can improve red blood cell production from stem cells. *Cell Stem Cell* 18, 73–78 (2016).
39. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34, 525–527 (2016).
40. Pimentel, H., Bray, N. L., Puente, S., Melsted, P. & Pachter, L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat. Methods* 14, 687–690 (2017).
41. Hahne, F. & Iraneck, R. Visualizing genomic data using Gviz and Bioconductor. *Methods Mol. Biol.* 1418, 335–351 (2016).
42. Ulirsch, J. C. et al. Altered chromatin occupancy of master regulators underlies evolutionary divergence in the transcriptional landscape of erythroid differentiation. *PLoS Genet.* 10, e1004890 (2014).
43. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578 (2012).
44. Xue, S. et al. RNA regulons in Hox 5′ UTRs confer ribosome specificity to gene regulation. *Nature* 517, 33–38 (2015).
45. Koontz, L. TCA precipitation. *Methods Enzymol.* 541, 3–10 (2014).
46. Engebrecht, J. M. et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341, 1237973 (2013).
47. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2013).
48. Li, W., Wang, W., Uren, P. J., Penalva, L. O. F. & Smith, A. D. Riborex: fast and flexible identification of differential translation from Ribo-seq data. *Bioinformatics* 33, 1735–1737 (2017).
49. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26 (2011).
50. Klass, D. M. et al. Quantitative proteomic analysis reveals concurrent RNA-protein interactions and identifies new RNA-binding proteins in *Saccharomyces cerevisiae*. *Genome Res.* 23, 1028–1038 (2013).
51. Simsek, D. et al. The mammalian ribo-interactome reveals ribosome functional diversity and heterogeneity. *Cell* 169, 1051–1065 e18 (2017).

**Acknowledgements**
We are grateful to D. Nathan, G. Daley, L. Zon and members of the Sankaran laboratory for valuable guidance and suggestions. We are grateful to D. Tenen for assistance with RNA immunoprecipitation and H. Keshishian for assistance with mass spectra data access. This work was supported by National Institutes of Health grants nos. U01 HL117720, R01 DK103794 and R33 HL120791 (to V.G.S.) and P01 DK32094 (to N.M.), a gift from the Lodish Family to Boston Children’s Hospital (to V.G.S.) and the New York Stem Cell Foundation (to V.G.S.). V.G.S. is a NYSCF–Robertson Investigator.

**Author contributions**
A.B. and V.G.S. conceived and designed the study. A.B., M.M. and K.E.M. performed experiments and analyzed data. C.A.L. and J.C.U. performed analyses. C.R.H., M.S., J.L., Y.W., Y.H., X.W., L.G., C.M.R., X.A., H.A.C., N.M., S.A.C., J.-J.C., S.H.O. and E.S.L. provided experimental assistance, reagents and advice. V.G.S. supervised all experimental and analytic aspects of this work. N.M. and V.G.S. acquired funding. A.B., M.M. and V.G.S. wrote the manuscript with input from all authors.

**Competing interests**
The authors declare no competing interests.

**Additional information**
Extended data is available for this paper at https://doi.org/10.1038/s41588-019-0568-7. Supplementary information is available for this paper at https://doi.org/10.1038/s41588-019-0568-7. Correspondence and requests for materials should be addressed to A.B. or V.G.S. Reprints and permissions information is available at www.nature.com/reprints.
Extended Data Fig. 1 | See next page for caption.
Extended Data Fig. 1 | BCL11A protein and mRNA expression in fetal, newborn and adult erythroid cells. **a**, Representative flow cytometry plots showing CD71 and CD235a surface expression in newborn (left) and adult (right) at differentiation day 7. Mean ± s.d shown. \( n = 3 \); 3 biologically independent experiments. **b**, Representative westerns showing BCL11A expression, with GAPDH as control, in newborn (left) and adult (right) at differentiation day 7 (5 independent experiments). **c**, BCL11A mRNA expression (normalized to GAPDH), in newborn and adult at differentiation day 7 \( n = 3 \); 3 biologically independent experiments). Mean ± s.d shown. Two-sided Student t-test used. N.S., not significant; \( P = 0.7153 \). **d**, Representative westerns showing BCL11A expression, with GAPDH used as control, at differentiation days 4, 7, & 10 (5 independent experiments). **e**, BCL11A mRNA expression (normalized to GAPDH) in newborn and adult \( n = 3 \); 3 independent experiments) at differentiation days 4, 7, and 10. Mean ± s.d shown. Two-sided Student t-test used. N.S., not significant; \( P = 0.4395 \) (d4), \( P = 0.3051 \) (d7), \( P = 0.3672 \) (d10). **f**, XL isoform of BCL11A mRNA with 4 exons and qRT–PCR primers. FP, forward primer; RP, reverse primer. **g**, BCL11A mRNA expression (normalized to GAPDH), in newborn and adult \( n = 3 \); 3 independent experiments) with 2 independent primer sets at differentiation day 7. Mean ± s.d shown. Two-sided Student t-test used. N.S., not significant, \( P = 0.2365 \) (pair 1), \( P = 0.4099 \) (pair 2). **h**, Stacked bar graphs showing fetal (HbF, red) and adult (HbA, grey) hemoglobin abundance (by HPLC) in newborn and adult on differentiation day 16. **i**, Representative westerns showing BCL11A expression with GAPDH as control at differentiation days 4, 7, 10, and 12 in fetal and adult (3 independent experiments). Blots have been cropped and the corresponding full blots are available in the Source Data files.
Extended Data Fig. 2 | Assessment of BCL11A and other mRNAs expressed in newborn and adult erythroid cells using RNA sequencing. a, Scatter plots of gene expression, determined from RNA-sequencing reads, expressed as log2 TPM (transcripts per million reads) in adult and newborn primary erythroid cells. ‘r’ represents the Pearson correlation coefficient. b, Expression of BCL11A, represented as log2 TPM, between newborn (n = 2) and adult (n = 2) erythroid cells. Error bars show s.d. c, BCL11A mRNA structure and splicing is comparable between developmental stages. Sashimi plots depicting exon-exon spanning reads are shown for annotated isoforms of BCL11A. d, Representation of known BCL11A isoforms. e, Relative abundances of BCL11A isoforms in newborn (n = 2) and adult (n = 2). Error bars show s.d. No transcript was differentially expressed at P < 0.01 between the newborn and adult cells.
Extended Data Fig. 3 | BCL11A protein expression is regulated via translation by polysome-associated LIN28B. a, Western blots for BCL11A in the input and flow-through (FT) fractions of the immunoprecipitate in newborn (left) and adult (right) erythroid cells (2 independent experiments). b, Western blots for BCL11A and GATA1 in the total immunoprecipitate (IP) in newborn (left) and adult (right) erythroid cells after L-azidohomoalanine (L-AHA) labeling for 6 hours at day 7 of differentiation, followed by immunoprecipitation with BCL11A and GATA1 antibodies. c, Quantification of adult-BCL11A (blue) and newborn-BCL11A (purple) mRNAs across the different sucrose gradient fractions are shown as a percentage of the gradient. Cells were differentiated until day 7. Blots have been cropped and the corresponding full blots are available in the Source Data files.
Extended Data Fig. 4 | LIN28B association with polysomes and expression in newborn and adult cells. Representative LIN28B occupancy across polysome fractions in newborn erythroid cells at day 7 of differentiation. LIN28B abundance is probed by western blot, using RPL5 and RPS20 as controls. Please note two gaps in the western blot between sequential polysome fractions that were placed to avoid overloading of proteins (2 independent experiments). Blots have been cropped and the corresponding full blots are available in the Source Data files.
Extended Data Fig. 5 | LIN28B partially dissociates from polysomes after RNase A treatment. a, LIN28B occupancy across polysome fractions in erythroid cells either untreated (blue) or treated with RNase A (red). Experiment repeated 3 times. b, LIN28B abundance probed by western blot, using RPL5 and RPS20 as controls in the untreated sucrose gradient fractions. c, LIN28B abundance in polysome fractions with lysates digested with RNase A. Blots have been cropped and the corresponding full blots are available in the Source Data files.
Extended Data Fig. 6 | LIN28B expression in newborn and adult cells. a, Volcano plot of differentially expressed genes between adult (n = 2) and newborn (n = 2) erythroid cells. Each dot is a gene with the value of the β coefficient (x-axis) from the sleuth linear model and the corresponding measure of statistical significance (y-axis). LIN28B is the most over-expressed gene in newborn cells compared to adult. Statistical test: generalized linear model from sleuth. b, Expression of LIN28B, represented as log2 TPM, between newborn and adult erythroid cells. Error bars show s.d. ***P < 0.001. Statistical test: generalized linear model from sleuth.
Extended Data Fig. 7 | Effects of high-level LIN28B expression in adult erythroid cells. a, BCL11A mRNA levels (normalized to GAPDH expression), upon high-level LIN28B expression (GFP-high) in adult erythroid cells, assessed on differentiation day 7 (n = 3; 3 biologically independent experiments). Mean is plotted and error bars show s.d. Two-sided Student t-test used. **P < 0.01. b, Relative γ-globin expression as a percentage of total globins (γ- and β-globins), upon high-level LIN28B expression (GFP-high) in adult erythroid cells, on differentiation day 15 (n=3; 3 independent experiments). Error bars show s.d. ***P < 0.001. c, Representative flow cytometry plots showing CD71 and CD235a surface expression in control (left) and physiological level LIN28B expressing (right) adult erythroid cells at day 7 of differentiation. Data represents mean ± s.d., n = 3; 3 biologically independent experiments.
Extended Data Fig. 8 | See next page for caption.
Extended Data Fig. 8 | LIN28B associates with BCL11A mRNA. 

**a**, RNA-immunoprecipitation (RNA-IP) in newborn erythroid cells at day 7 of differentiation, with antibodies against LIN28B (4196, Cell Signaling) or control IgG. Detection of BCL11A, HMGA1, GATA1, ALAS2, LDB1, KLF1, and LMO2 mRNAs (n = 3; 3 independent experiments). Mean is plotted and error bars show s.d. Two-sided Student t-test used. **P < 0.01; N.S., not significant, P = 0.6395 (GATA1), P = 0.8782 (ALAS2), P = 0.8999 (LDB1), P = 0.9571 (KLF1), P = 0.9550 (LMO2).**

**b**, Detection of BCL11A, HMGA1, GATA1, ALAS2, LDB1, KLF1, and LMO2 mRNAs (n = 3; 2 independent experiments) in the input for LIN28B antibody (4196, Cell Signaling). Mean is plotted and error bars show s.d. Two-sided Student t-test used. N.S., not significant, P = 0.3867 (BCL11A), P = 0.8050 (HMGA1), P = 0.6420 (GATA1), P = 0.8656 (ALAS2), P = 0.6413 (LDB1), P = 0.3257 (KLF1), P = 0.7152 (LMO2).

**c**, RNA-IP in newborn erythroid cells with antibodies against LIN28B (A303-588A, Bethyl Labs) or control rabbit IgG. Detection of BCL11A (left), HMGA1 (center), and GATA1 (right) mRNAs (n = 3; 2 independent experiments). Mean is plotted and error bars show s.d. Two-sided Student t-test used. *P < 0.05; ***P < 0.001; N.S., not significant, P = 0.5393.

**d**, Detection of BCL11A (left), HMGA1 (center) and GATA1 (right) mRNAs (n = 3; 3 independent experiments) in the input for LIN28B antibody (A303-588A, Bethyl Labs). Mean is plotted and error bars show s.d. Two-sided Student t-test used. N.S., not significant, P = 0.5261 (BCL11A), P = 0.4871 (HMGA1), P = 0.9464 (GATA1).
Extended Data Fig. 9 | CLIP-seq of LIN28B in newborn erythroid cells identifies genome-wide binding peaks. a, Genomic annotation of LIN28B binding sites. Peaks passing a 1% IDR for the consensus between replicates and a 5% false-discovery rate for each replicate are shown. b, 4-mer motifs associated with LIN28B binding sites. The dashed lines represent a threshold of 60%. 17 4-mers are present at >60% in both exons and UTRs, including GGAG, GAAG, and AAGA. c, Fold enrichment of genomic distance covered by LIN28B binding peaks compared to the genome-wide proportions. d, Rank-order enrichment of LIN28B binding peaks genome-wide. Each peak’s enrichment, measured by its -log10 q-value is plotted against its rank genome-wide for both replicates (n = 2). Statistical test: macs2 peak calling algorithm. e, Total coverage and mutation proportion for each replicate of the LIN28B/BCL11A binding site. f, Volcano plots showing the log-fold change with Benjamini-Hochberg adjusted P-values for LIN28B target genes compared between newborn (n = 2) and adult (n = 2) proerythroblasts. Statistical test: generalized linear model from sleuth.
Extended Data Fig. 10 | Suppression of γ-globin by BCL11A expression in newborn erythroid cells. a, γ-globin levels upon BCL11A expression in newborn erythroid cells on differentiation day 12 (n = 3; 3 biologically independent experiments) in control and BCL11A expressing cells. Mean is plotted and error bars show s.d. Two-sided student t-test used. ****P < 0.0001. b, Representative western blots showing BCL11A expression from lentiviral construct in newborn erythroid cells at day 12 of differentiation. GAPDH is used as a loading control (3 independent experiments). c, Representative flow cytometry plots showing CD71 and CD235a surface expression in control (left) and BCL11A expressing (right) newborn erythroid cells at days 8, 10 and 12 of differentiation (3 biologically independent experiments). Blots have been cropped and the corresponding full blots are available in the Source Data files.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- n/a - Confirmed
- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
- Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Differential gene expression analyses were conducted using transcript abundance estimation from kallisto v0.44 and sleuth v0.30. Peak calling was performed using MACS2 version 20160309.

Data analysis

Data were analysed using Microsoft Excel (version 16.16), GraphPad Prism 7, FlowJo v 10.3, Image J v1.8.0, R version 3.4.3, TopHat v2.1.0 and BioConductor v1.7.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The accession code for CLIP, RNA and Ribosome sequencing data is GSE118359.
The following figures have associated raw data: Figs. 1i,1j, 2a,2b,4a,4b and S2, S4b, S4c, S7
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences  ☐ Behavioural & social sciences  ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | No sample size power calculations were performed. Since independent human donors are used, 3 independent biological replicates were generally used for most experiments. |
|------------|--------------------------------------------------------------------------------------------------|
| Data exclusions | Data were not excluded from analysis. |
| Replication | To ensure the reproducibility of experimental findings, replicates were carried out. Experiments involving human erythroid progenitors were repeated with multiple different donors to ensure consistency. |
| Randomization | No randomization was performed for this study. |
| Blinding | No blinding was performed for this study. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

| Materials & experimental systems | Methods |
|-------------------------------|---------|
| n/a | Involved in the study |
| ☐ | Antibodies |
| ☒ | Eukaryotic cell lines |
| ☐ | Palaeontology |
| ☒ | Animals and other organisms |
| ☒ | Human research participants |
| ☐ | Clinical data |
| ☒ | ChIP-seq |
| ☒ | Flow cytometry |
| ☒ | MRI-based neuroimaging |

Antibodies

Antibodies used

Details (supplier name, catalog number, clone name, lot number, dilution) of all antibodies used are listed in Supplementary Table 3

Validation

All antibodies in this study were used according to manufacturer’s instructions and dilutions.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

ATCC. Cell lines used: 293T, K562.

Authentication

Identity of cell lines were validated by STR analysis.

Mycoplasma contamination

Mycoplasma testing was routinely performed on all cells used in the study, and confirmed to test negative.

Commonly misidentified lines (See ICLAC register)

None of the cell lines used are listed in the ICLAC database.
Flow Cytometry

Plots

Confirm that:
- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a `group` is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

| Sample preparation | Cell cultures were harvested by centrifugation (500g, 5min), washed with phosphate-buffered saline (PBS, Invitrogen:14190-250) and stained with anti-human antibodies on ice for 20 min. All antibodies used are listed in Supplementary Table 3. Propidium iodide (PI) or 4,6-diamidino-2-phenylindole (DAPI) were used for live/dead cell discrimination. |
|--------------------|--------------------------------------------------------------------------------------------------|
| Instrument         | Flow cytometry analyses were carried out on Becton Dickinson (BD) LSRII, LSR Fortessa (BD), or Accuri C6 (BD) instruments. |
| Software           | All data was analyzed using FlowJo software (version 10.3). |
| Cell population abundance | Cell sorting, and thus the purity of the samples were carried out based on green fluorescent protein (GFP) or red fluorescent protein (RFP) expression. 0.3-5 million post-sorted samples were collected, depending on the experiment. |
| Gating strategy    | Debris were initially removed with the FSC/SSC gating strategy. Live/dead discrimination was then carried out with DAPI/PI stains. Where applicable, these cells were then assessed for the expression of CD71 and CD235a surface markers, and gated based on the unstained control used in the experiment. Where cell sorting was applied, stage-matched unstained cells were used to demarcate the boundaries between "negative" and the "positive" cells. |

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.