Published online 17 January 2014 *Nucleic Acids Research, 2014, Vol. 42, No. 6* ***4031–4042***
doi:10.1093/nar/gkt1388

NAR Breakthrough Article

A dynamic alternative splicing program regulates gene expression during terminal erythropoiesis

Harold Pimentel[1], Marilyn Parra[2], Sherry Gee[2], Dana Ghanem[2], Xiuli An[3], Jie Li[3], Narla Mohandas[3], Lior Pachter[1,4,5] and John G. Conboy[2,*]

[1]Department of Computer Science, University of California, Berkeley, CA 94720, USA, [2]Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA, [3]Red Cell Physiology Laboratory, New York Blood Center, New York, NY 10065, USA, [4]Department of Mathematics, University of California, Berkeley, CA 94720, USA and [5]Department of Molecular & Cell Biology, University of California, Berkeley, CA 94720, USA

Received October 29, 2013; Revised December 16, 2013; Accepted December 17, 2013

ABSTRACT

Alternative pre-messenger RNA splicing remodels the human transcriptome in a spatiotemporal manner during normal development and differentiation. Here we explored the landscape of transcript diversity in the erythroid lineage by RNA-seq analysis of five highly purified populations of morphologically distinct human erythroblasts, representing the last four cell divisions before enucleation. In this unique differentiation system, we found evidence of an extensive and dynamic alternative splicing program encompassing genes with many diverse functions. Alternative splicing was particularly enriched in genes controlling cell cycle, organelle organization, chromatin function and RNA processing. Many alternative exons exhibited differentiation-associated switches in splicing efficiency, mostly in late-stage polychromatophilic and orthochromatophilic erythroblasts, in concert with extensive cellular remodeling that precedes enucleation. A subset of alternative splicing switches introduces premature translation termination codons into selected transcripts in a differentiation stage-specific manner, supporting the hypothesis that alternative splicing-coupled nonsense-mediated decay contributes to regulation of erythroid-expressed genes as a novel part of the overall differentiation program. We conclude that a highly dynamic alternative splicing program in terminally differentiating erythroblasts plays a major role in regulating gene expression to ensure synthesis of appropriate proteome at each stage as the cells remodel in preparation for production of mature red cells.

INTRODUCTION

Alternative pre-messenger RNA (mRNA) splicing enables individual genes to generate multiple protein products that differ structurally and functionally by insertion or deletion of important functional domains encoded by alternative exons. During normal development and differentiation, dynamic changes in the expression or activity of the splicing regulatory machinery coordinately modulate networks of alternative splicing events in a spatiotemporal manner. Post-transcriptional RNA processing can thereby modulate essential protein functions according to the physiological requirements of the cell by regulating coherent biological pathways (1). Conversely, network perturbations caused by mutations that alter splicing factor expression or activity underlie an array of complex human diseases (2). The experimental work supporting these concepts has been performed primarily in non-hematologic cell types. However, recent studies have revealed that normal T cells execute a complex splicing program (3) and that splicing factor mutations are associated with hematological cancers, including myelodysplasia (4–6), demonstrating the importance of alternative splicing in hematology. Characterization of the alternative splicing program in human erythroblasts undergoing terminal erythroid differentiation will reveal novel insights into RNA regulatory pathways that drive cell differentiation and provide a basis for identifying splicing defect in human erythroid diseases.

Alternative isoforms of various erythroid transcripts have been reported, and a modest number of differentiation stage-specific switches in alternative splicing patterns are known (7). Best studied mechanistically is the upregulation of splicing efficiency for protein 4.1R alternative exon 16, which encodes part of the spectrin–actin binding domain required for optimal assembly of a mechanically competent membrane skeleton (8,9), an essential structure for mature erythrocyte function. This exon is

*To whom correspondence should be addressed. Tel: +1 510 486 6973; Fax: +1 510 486 6746; Email: JGConboy@lbl.gov

predominantly skipped in early erythroblasts but included efficiently in late erythropoiesis (10,11), and several RNA binding proteins/splicing factors that influence exon 16 splicing have been identified (12–15). Underscoring the importance of this splicing switch, failure to include exon 16 causes mechanically unstable red cells and aberrant elliptocytic phenotype with anemia (16,17).

To comprehensively explore the alternative splicing landscape during terminal erythroid differentiation, we used an RNA-seq strategy to analyze and compare transcriptomes of highly purified human erythroblasts cultured *in vitro* from $CD34^+$ cord blood progenitors (18). RNA-seq allows a robust high-resolution assessment of the transcriptome, but there remain computational challenges in data interpretation. Using extensions of current transcript abundance estimation tools combined with non-parametric statistical methods, we found an extensive alternative splicing program that is significantly enriched in genes controlling cell cycle, organelle organization, chromatin function and RNA processing. Importantly, hundreds of these alternative splicing events are regulated in a differentiation stage-specific manner, with most switches in exon inclusion/exclusion efficiency occurring in late-stage erythroblasts concurrent with extensive cellular remodeling as erythroblasts transition from a highly proliferative state to a terminally differentiated state. Finally, we discovered a subset of splicing switches that introduce premature translation termination codons (PTCs), thus decreasing the proportion of full-length coding mRNAs and downregulating gene expression via induction of nonsense-mediated decay (NMD). Post-transcriptional RNA processing pathways may be regulated by this mechanism in late erythroid differentiation, as numerous RNA binding proteins exhibit elevated expression of PTC exons in the most mature erythroblasts. These studies support the central hypothesis that human erythroblasts execute a surprisingly complex, differentiation stage-specific alternative splicing program that is essential for normal proliferation and differentiation.

MATERIALS AND METHODS

Human erythroblast populations

CD34+ cells were enriched from cord blood, cultured under conditions that promote erythroblast differentiation, and further enriched for stage-specific erythroblast populations by fluorescence-activated cell sorting (FACS) exactly as described (18). By gating narrow windows, high cell purity is attained for five discrete erythroblast populations (>90%), representing the last four cell divisions in terminal erythropoiesis. In some experiments, the cell populations enriched for proerythroblasts (proE) (culture day 9) and orthochromatic erythroblasts (orthoEs) (culture day 16) were treated with 100 μg/ml cycloheximide for 4 h to inhibit nonsense-mediated decay of RNA transcripts containing PTCs.

RNA-seq analysis

Total RNA was extracted using an RNeasy Plus Mini Kit (QIAGEN) from sorted human erythroblasts at distinct stages of erythroid maturation. The integrity of the RNA was evaluated on an Agilent Bioanalyzer, and samples with RNA integrity numbers (RIN) >8.0 were prepared for sequencing. Poly(A)-tailed RNA was subsequently prepared by the Epigenomics Core, Cornell Medical College, using the mRNA Seq Sample Prep Kit (Illumina Inc., San Diego, CA, USA) and used to create libraries for HiSeq2000 sequencing (Illumina).

Hundreds of millions of 50-nt RNA-seq reads were obtained from each of the 15 erythroblast samples, representing five distinct stages of differentiation from three independent sorting experiments. RNA-seq fragments were aligned to the Ensembl-annotated transcriptome version 67 (19) [hg19 human genome build GRCh37 (20)] using the Bowtie aligner (21). Read statistics are given in Supplementary Table S1. Transcript-level estimates were obtained using the transcript abundance estimation tool, eXpress (22), and were expressed as fragments per kilobase of exon per million fragments mapped (FPKM) as described previously (23,24). For each exon of every Ensembl gene, transcripts from the gene were divided into two sets: the set of isoforms including the exon (termed inclusion) and the set of isoforms with an intronic region spanning it (termed overlap). Expression of individual exons in 'exon-inclusion' isoforms relative to the total expression of all isoforms was represented as percent spliced in (PSI) and computed by calculating the ratio FPKM (inclusion)/FPKM (inclusion, overlap). Because only ratios of reads were compared, a library size normalization step between samples was not needed.

To classify exons as differentially expressed between the proE and orthoE samples, we used the estimated PSI values with the *samr R* package (25), assuming continuous data. This resulted in a test similar to a *t*-test, but where the null distribution was generated by a permutation test non-parametrically from the data. In the same manner, we estimated the PSI of each exon in the basophilic (basoE) and polychromatophilic (polyE) erythroblast populations. We then considered any exon for which the inclusion plus overlap FPKM was in the lower quartile to be insufficiently expressed and discarded it from the analysis. We also removed any exon for which a contributing isoform had a contradiction in coverage by having long spanning regions of 0 coverage (see Supplementary Methods). From this list filtered for low gene expression, we also filtered out exons with minimal alternative splicing (PSI <0.05 or >0.95) at any differentiation stage, because empirically we found that this PSI filter greatly reduced the incidence of false positives. Data for the PSI-filtered set of splicing events are shown in Supplementary Table S2, and the larger unfiltered PSI data set is provided in Supplementary Table S3. The filtered list of alternatively spliced exons was used as input into GOrilla (*Gene Ontology enRIchment anaLysis and visuaLizAtion tool*) for Gene Ontology (GO) analysis (26) (Table 1) and for analyzing stage-specificity of splicing switches (Figure 5).

Besides the computationally defined alternative splicing switches, a few additional examples were discovered in genes of interest by manual inspection of RNA-seq reads mapped to the human genome. Exons that did not

Table 1. GO terms enriched in alternatively spliced erythroblast genes

Rank	GO term	Description	*P*-value	FDR q-value
1	GO:0044260	Cellular macromolecule metabolic process	1.74E-11	1.57E-07
2	GO:0006996	Organelle organization	1.73E-10	7.84E-07
3	GO:0022402	Cell cycle process	4.01E-09	1.21E-05
6	GO:0016043	Cellular component organization	5.48E-08	8.27E-05
8	GO:0016568	Chromatin modification	1.41E-07	1.60E-04
10	GO:0006325	Chromatin organization	1.61E-07	1.46E-04
12	GO:0007049	Cell cycle	5.63E-07	4.25E-04
13	GO:0080090	Regulation of primary metabolic process	7.68E-07	5.35E-04
14	GO:0010468	Regulation of gene expression	8.07E-07	5.22E-04
15	GO:0044265	Cellular macromolecule catabolic process	9.00E-07	5.43E-04
18	GO:0051276	Chromosome organization	1.37E-06	6.90E-04
19	GO:0000278	Mitotic cell cycle	1.72E-06	8.21E-04
20	GO:0016569	Covalent chromatin modification	2.59E-06	1.17E-03
21	GO:0016570	Histone modification	2.75E-06	1.19E-03
22	GO:0090304	Nucleic acid metabolic process	2.97E-06	1.22E-03
23	GO:0048285	Organelle fission	3.84E-06	1.51E-03
30	GO:0000279	M phase	7.27E-06	2.19E-03
33	GO:0010467	Gene expression	1.07E-05	2.94E-03
34	GO:0000086	G2/M transition of mitotic cell cycle	1.14E-05	3.03E-03
35	GO:0000226	Microtubule cytoskeleton organization	1.48E-05	3.81E-03
39	GO:0006511	Ubiquitin-dependent protein catabolic process	1.86E-05	4.32E-03
45	GO:0016070	RNA metabolic process	2.48E-05	4.98E-03
50	GO:0016573	Histone acetylation	2.86E-05	5.18E-03
56	GO:0000280	Nuclear division	5.01E-05	8.09E-03
57	GO:0007067	Mitosis	5.01E-05	7.95E-03
58	GO:0044772	Mitotic cell cycle phase transition	5.27E-05	8.22E-03
60	GO:0000090	Mitotic anaphase	5.67E-05	8.55E-03
62	GO:0051301	Cell division	8.95E-05	1.31E-02
67	GO:0070925	Organelle assembly	1.25E-04	1.69E-02
76	GO:0043484	Regulation of RNA splicing	1.90E-04	2.27E-02

computationally meet a ΔPSI threshold of 10% but showed strong evidence for stage-specific changes in the ratio of exon junction reads supporting inclusion versus exon exclusion reads were tested by reverse-transcriptase polymerase chain reaction (RT-PCR); examples of splicing switches detected in this manner are MEF2D, PICALM and MTMR3.

Validation of computationally predicted alternative exons

Many alternative splicing events were verified by manual inspection of mapped reads on the human genome browser, to filter out exons not supported by multiple exon-inclusion and exon-skipping reads. In selected cases, RT-PCR analysis was performed to visualize the amplified products corresponding to inclusion and skipping events. Total RNA was reverse-transcribed into cDNA using random primers and the Superscript III First Strand Synthesis System (Invitrogen), and PCR analysis was performed using AccuTaq polymerase (Sigma) using primers located in constitutive exons located upstream and downstream of each candidate alternative exon. Primer sequences are given in Supplementary Table S4. For initial RT-PCR validations, RNA from unsorted day 9 (proE-enriched) and day 16 (orthoE-enriched) cells was used; these results were subsequently confirmed with RNA purified from highly enriched (FACS-sorted) proEs and orthochromatophilic erythroblasts.

RESULTS

Differentiating erythroid cells execute a robust alternative splicing program

Highly purified human erythroblasts were obtained from *in vitro* cultures by FACS using differentiation stage-restricted cell surface markers as previously described (18). Stage-specific RNA-seq libraries were prepared from five morphologically distinct populations of successively more differentiated erythroblasts, each separated from the next by one cell division: proEs, early basophilic erythroblasts (e-basoE), late basophilic erythroblasts (l-basoE), polychromatophilic erythroblasts (polyE) and finally orthochromatophilic erythroblasts (orthoE), the last stage before enucleation (Figure 1). Three independent preparations of sorted cells were analyzed to provide biological replicates.

Alternative splicing events in differentiating erythroblasts were identified following the strategy shown in Figure 2. A total of 1.5 billion reads representing the transcriptomes of all 15 erythroblast samples were mapped to the Ensembl annotated transcriptome and analyzed using numerous tools to derive exon-level alternative splicing information as described in 'Methods' section. The initial set of 243 464 candidate alternative exons was filtered to remove low-expression genes, and exons with minimal alternative splicing, to derive a data set of 3784 alternative exons in 2250 genes (Supplementary Table S2). Validation of computationally predicted alternative

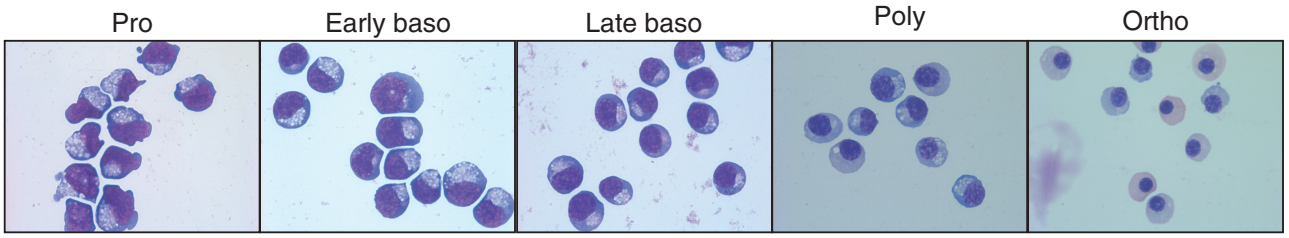

Figure 1. Morphology of differentiation stage-specific erythroblast populations purified by FACS as described previously (18).

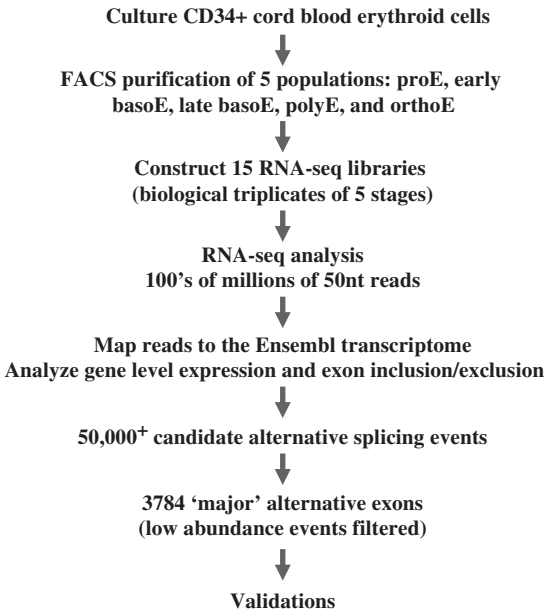

Figure 2. Flow chart showing the process by which major alternative splicing events were detected in discrete populations of erythroblasts at specific stages of their terminal differentiation.

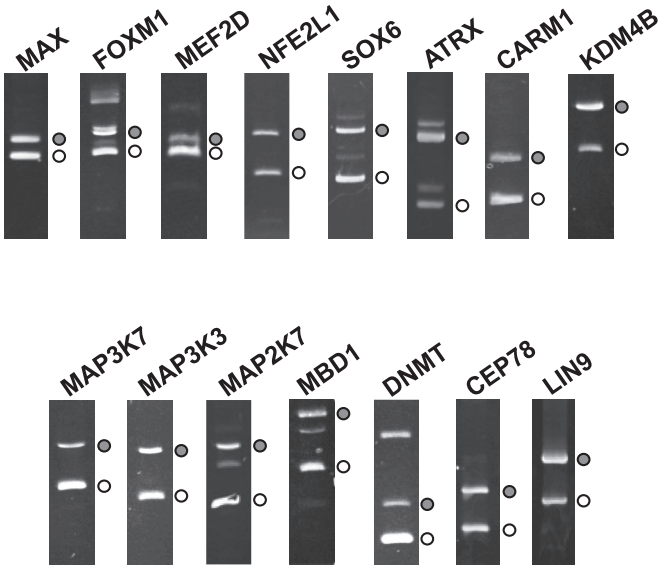

Figure 3. RT-PCR validation of alternative splicing events in human erythroblasts. Shown are amplification products containing (filled circles) or lacking (empty circles) alternative exons amplified from erythroblast transcripts. Gene names are indicated above each lane.

splicing events was performed by manual inspection of RNA-seq reads mapped to the human genome and by RT-PCR analysis of the relevant transcript regions. Figure 3 illustrates representative alternative splicing events in genes involved in transcriptional regulation (NFE2L1, MEF2D, SOX6, ATRX, FOXM1, MAX), DNA and histone modifications (DNMT1, MBD1, CARM1, KDM4B) and signal transduction pathways (MAP2K7, MAP3K3, MAP3K7). The majority of these alternative exons did not substantially alter their splicing efficiency between proEs and orthoEs; i.e. similar results were observed by computational and experimental analyses independent of differentiation stage. However, alternative exons in NFE2L1, MEF2D and CARM1 were among those observed to splice differently as the cells matured (discussed below in Figure 4). Of interest, SOX6 and ATRX are important regulators of globin gene expression (27–29). These results demonstrate that differentiating erythroblasts execute an extensive alternative splicing program that can regulate protein structure and function via control of pre-mRNA splicing; moreover, when the affected proteins are transcription factors, important secondary effects on transcript levels of other genes might amplify downstream cellular consequences.

To investigate more broadly which biological processes are impacted by the erythroid splicing program, we performed GO analysis. GO terms significantly enriched among genes expressing major alternative splicing events, relative to all genes expressed (above the lowest quartile) in erythroblasts, are shown in Table 1. The top terms included numerous descriptors associated with chromatin structure and function, cell cycle regulation, organelle organization and regulation of RNA splicing. Among the read-validated alternative splicing events, we found at least 24 differentially expressed exons in 22 genes encoding enzymes that participate in histone modifications (30) (Supplementary Table S5), and at least 65 high-confidence alternative splicing events were detected in 48 genes encoding RNA binding proteins (Supplementary Table S6). As indicated in the tables, many of these alternative splicing events are not annotated among Refseq transcripts, and ~40% are predicted to introduce PTCs that induce NMD. All of these splicing events passed low-expression and low-alternative splicing filters, and all were validated by RNA-seq read mapping. These observations suggest a high level of complexity in regulation of

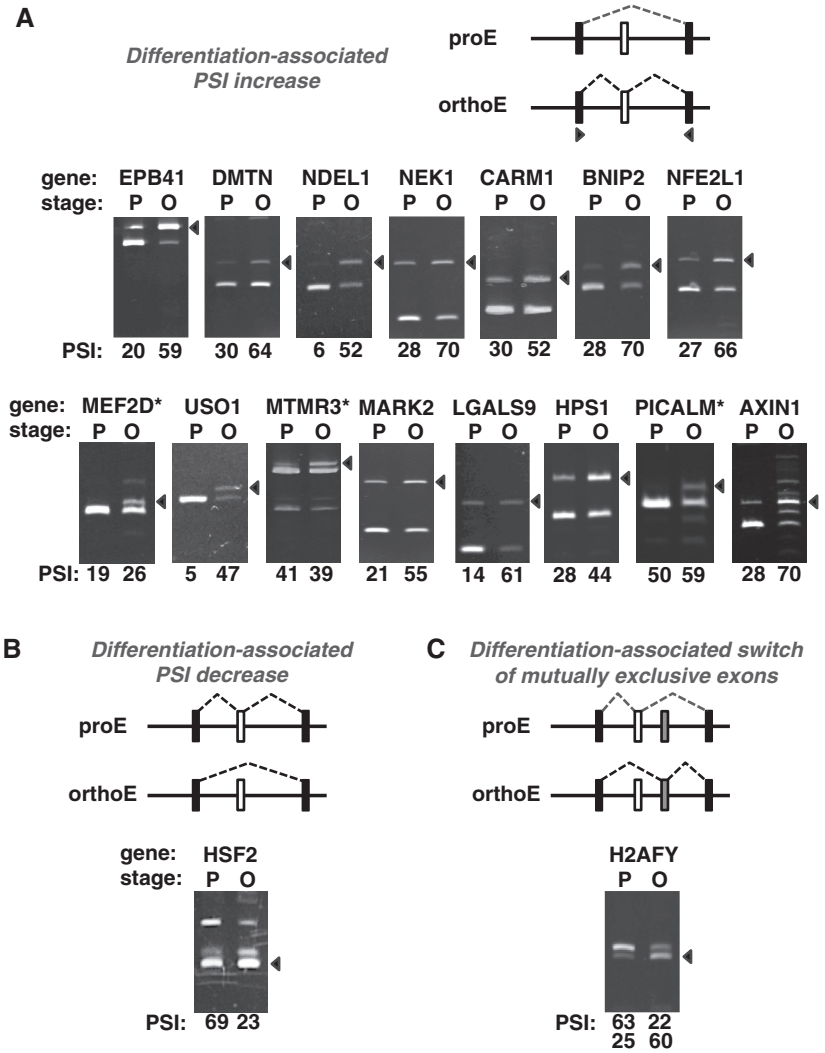

Figure 4. Differentiation stage-specific alternative splicing switches in human erythroblasts. (**A**) Alternative exons that exhibit stage-specific upregulation in splicing efficiency. Gels show splicing changes assessed by RT-PCR using primers located in flanking constitutive exons. Amplification products from proE RNA and orthoE RNA are shown in the left and right lanes, respectively. Gene name is indicated above each gel, while the calculated PSI is shown below each lane. Arrowheads indicate exon inclusion products. Asterisk indicates genes for which PCR results indicate larger splicing changes than were predicted bioinformatically. (**B**) Alternative splicing event that exhibits a differentiation-associated decrease in PSI as cell matures. (**C**) A differentiation-associated switch involving mutually exclusive exons of 100 and 91 nt, respectively. The size difference allows electrophoretic separation of the alternative products.

erythroid gene expression, and a strong potential that these splicing events have downstream effects on the abundance and structure of transcripts of numerous other genes. Thus, important biological pathways express a rich transcript diversity during late erythropoiesis.

Stage-specific switches in alternative splicing remodel the transcriptome during late erythroid differentiation

Alternative splicing switches can play key roles during differentiation and development by selective up- or downregulation of individual exons that alter function of the encoded protein(s). To determine the extent of splicing transitions during erythroid differentiation, we first compared PSI values for alternative exons in early (proE) and late (orthoE) erythroblast populations. In all, 1166 exons for which the difference in PSI values (ΔPSI) exceeded 10% were identified by this approach; 266 of these exhibited ΔPSI values exceeding 25%. A subset of splicing switches that alter protein coding domains was validated by RT-PCR using primers in the flanking constitutive exons (Figure 4). The previously described switch in protein 4.1R (EPB41) pre-mRNA splicing served as a positive control; as expected, splicing of exon 16 exhibited a substantial increase in PSI in orthoE cells compared with a much lower splicing efficiency in proE (Figure 4A). Similar to 4.1R, many of the newly discovered switches involve upregulation of splicing in orthoE (Figure 4A), but examples of downregulation (Figure 4B) and switching between mutually exclusive exons (Figure 4C) were also observed. For some of these, e.g. NDEL1, USO1, MEF2D and MTMR3, the 'new' isoforms with potentially new functions can be generated in orthoE

Table 2. Alternative splicing switches that alter coding domains

gene	Exon size	ΔPSI (proE to orthoE)	Gene function (biological process)
EPB41	63	+50	Cortical actin cytoskeleton
EPB49	66	+34	Actin cytoskeleton organization
NDEL1	35	+46	Centrosome localization
NEK1	132	+42	Cell division; mitosis
NUMA1	42	−18	Establishment of mitotic spindle orientation
BIRC5	69	+11	Cell division; cytokinesis
MARK2	162	+34	Regulation of cytoskeleton organization; protein serine/threonine kinase activity
MTMR3	27	−2[a]	Peptidyl-tyrosine dephosphorylation; phosphatidylinositol dephosphorylation
CARM1	69	+22	Methyltransferase activity
NFE2L1	90	+39	Sequence-specific DNA binding transcription factor activity
H2AFY	100[b]	−41	Chromatin modification
H2AFY	91[b]	+35	Chromatin modification
MEF2D	24	+7[a]	Regulation of transcription from RNA polymerase II promoter
PICALM	24	+9[a]	Clathrin coat assembly; hemopoiesis
BNIP2	36	+27	Apoptotic process
USO1	21	+42	Intracellular protein transport
HPS1	99	+16	Lysosome organization
LGALS9	96	+47	Carbohydrate binding

[a]Read map data and RT-PCR indicate stronger switch than predicted by computational analysis.
[b]Mutually exclusive exons.

that were absent or poorly expressed in proE. Except for USO1, PICALM1 and LGALS9, most of the genes with upregulated coding exons were also upregulated at the gene expression level, relative to global gene expression levels, further supporting the idea that the new isoforms might play important functions in late erythroblasts.

Consistent with a role for alternative splicing in important erythroid pathways, several validated switches in alternative splicing map to genes that function in cytoskeletal organization, cell division and chromatin function (Table 2). One of the biggest switches occurred in the NDEL1 (nuclear distribution factor E-homolog-like1) gene that is important for nuclear migration and nucleokinesis (31), while a less dramatic switch was predicted in survivin (BIRC5), a cytokinesis factor with a novel role in erythroblast enucleation (32,33). In the H2AFY gene, relative expression of mutually exclusive 100- and 91-nt exons reversed almost completely between proE and orthoE (Figure 4A). This switch alters the structure and function of the modified histone, macroH2A1, which can replace the canonical H2A to generate functional differences in chromatin activity (34–36). Consistent with the H2AFY splicing switch in terminal erythroid differentiation, differences in the relative abundance of the two isoforms have been reported in other tissues rich in proliferating versus differentiating cells (37). Other switches of likely functional importance occurred in kinases having roles in chromosome segregation and cytokinesis [NEK1, (38–40)] and microtubule- and actin-based cytoskeletal networks [MARK2, (41)] in a clathrin assembly gene (PICALM) associated with anemia and iron metabolism abnormalities in mouse mutants (42,43); and the mitotic assembly protein NUMA1.

To explore the hypothesis that alternative splicing networks can impact erythroblast remodeling, we examined the differentiation stage-specificity of splicing transitions. PSI as a function of erythroblast stage is shown in Figure 5A for several PCR-validated alternative exons that exhibited substantially increased inclusion between proE and orthoE (ΔPSI values $\geq$34). For these genes, most of the change in alternative splicing occurred in the last two (most mature) erythroblast populations, although lesser transitions in splicing efficiency occurred at earlier stages. Extending the analysis to the broader set of alternative exons, we next determined ΔPSI values across each of the four-stage transitions: proE to early basoE, early to late basoE, late basoE to polyE and polyE to orthoE. Figure 5B shows that ~155 stage-specific splicing switches with $\Delta PSI > 25$ were identified in these cultures, and the great majority of these occurred during the last two cell divisions. Of note, this is a conservative estimate that counts only exons having PSI values between 5 and 95 in all 15 samples assayed. These results demonstrate that most splicing transitions occur in late erythroblasts, temporally correlated with extensive cellular remodeling as the cells prepare to enucleate, and supporting the hypothesis that splicing changes play a determinative role in promoting late erythroid differentiation.

Alternative splicing coupled to NMD regulates expression of selected RNA processing factors

Alternative splicing coupled to NMD (AS-NMD) can regulate gene expression by splicing-mediated introduction of PTCs that induce nonsense-mediated degradation of the affected transcripts. PTC-containing transcripts can be generated by inclusion of exons that encode stop codons or alter translational reading frame (PTC-upon-exon inclusion), or by exon skipping events that alter reading frame (PTC-upon-exon skipping). Having noted that a number of RNA binding proteins express PTC transcripts in erythroblasts (Supplementary Table S4), we wondered whether AS-NMD was functioning to maintain

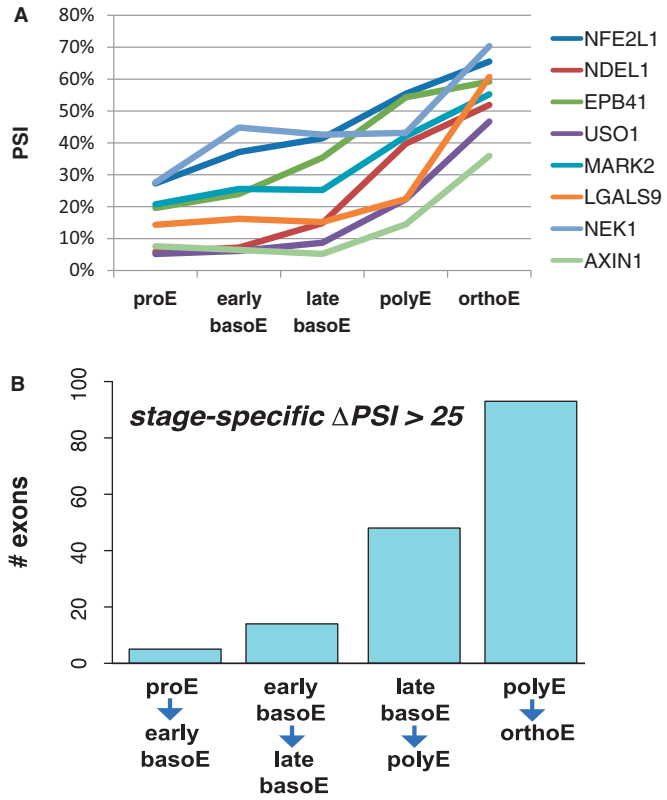

Figure 5. Stage-specificity of splicing switches during erythroblast differentiation. (**A**) Splicing efficiency of selected splicing switches across erythroblast populations. (**B**) Broader analysis of alternative exons exhibiting a splicing switch during erythroblast maturation. Number of exons with $|\Delta PSI| > 25$ is represented by bar height at the indicated stage transitions.

homeostatic control of expression as described in previous studies (44–46). Instead, PSI and gene expression data indicated a different situation characterized by substantial stage-specific changes in relative expression of PTC transcripts (Figure 6). Increased proportions of PTC-containing transcript isoforms were expressed in late erythroblasts by either increases in exon inclusion or by increases in exon skipping (Figure 6A). Many of the predicted increases are substantial; e.g. the SNRNP70 gene switches from PSI = 13 to PSI = 43 for the smaller PTC-upon-inclusion exon. Similarly, well-known AS-NMD targets in the SR protein family including SRSF2, SRSF3, SRSF6 and SRSF7 also showed large ΔPSI switches from proE to orthoE (ΔPSI values of +12 to +47). Other classes of AS-NMD candidates include hnRNP proteins (D, L, LL and M), small nuclear ribonucleic particle proteins (SNRPA1, SNRNP70 and U2AF1), alternative splicing regulators [TRA2A (45), TRA2B] and a 3' processing factor (CPSF4). In a few cases, an increased proportion of PTC transcripts in late erythroblasts can occur in the context of PTC-upon-exon skipping events (Figure 6A, lower panel). SNRPA1, for example, expressed mainly full-length transcripts in proE as indicated by PSI = 94, but the decrease to PSI = 53 in orthoE predicts substantial expression of PTC isoforms in orthoE. However, it is important to note that some known [SRSF11, SNRPB, (45)] or candidate (TIAL1 and HNRNPD) PTC splicing events showed no evidence of stage-specific regulation (Figure 6B, upper panel), and a few (RHD, BCL2L12, CLK1 and CLK4) were regulated so as to decrease the proportion of PTC isoforms in orthoE (Figure 6B, lower panel). RT-PCR validation of splicing switches involving PTC exons enriched in orthoE or proE is provided in Figure 7. We propose that splicing-mediated generation of non-productive (PTC-containing) transcripts is an important post-transcriptional regulatory mechanism for controlling a selected subset of RNA processing factors in late-stage erythroblasts.

In principle, elevated expression of PTC transcripts could be due to increased production (splicing regulation) or to decreased degradation (NMD regulation). We reasoned that simple loss of NMD activity could not explain maturation-associated increases in PTC exons, as only selected PTC exons were upregulated in orthoE and a few were actually downregulated in these cells. This hypothesis was investigated experimentally by assaying whether inhibition of NMD could affect steady-state levels of PTC exons. If NMD is active in normal orthoE, then steady-state PSI values for PTC exons should increase due to selective stabilization of these transcripts in cycloheximide-treated erythroblasts inhibited for NMD. Figure 8 shows that PTC-upon-inclusion exons in SNRNP70 and HNRPLL had substantially elevated PSI values in NMD-inhibited cells, with PSI of a 72-nt exon in SNRNP70 increasing moderately in proE-enriched day 9 erythroblasts and dramatically in orthoE-enriched day 16 erythroblasts (upper panel). NMD inhibition also increased the proportion of PTC transcripts from the BCL1L12 and CLK4 genes that were generated by exon skipping and were more enriched in proEs (Figure 8, lower panel). We conclude that the NMD machinery is still active in late erythroblasts and therefore that the accumulation of PTC isoforms in these normal cells is not primarily due to loss of NMD, but instead must be due, at least in part, to regulation at the level of alternative splicing. Moreover, as the 'missing' quantity of PTC isoforms in normal cells must have been degraded, we propose that alternative splicing coupled to NMD (AS-NMD) is a mechanism for downregulating gene expression in late erythroblasts, analogous to the intron retention mechanism used by late-stage granulocytes to downregulate selected genes (47). Consistent with this, 9 of 11 RBP genes with upregulated PTC events (82%) were downregulated at the gene expression level relative to total mRNAs in late erythroblasts. In contrast, among genes shown in Figure 4 with validated upregulation of non-PTC exons in orthoE, 12 of 15 genes (80%) exhibited increases in gene expression level.

DISCUSSION

The erythroblast system provides a unique opportunity to study the role of post-transcriptional RNA processing in a differentiation stage-specific manner during the extensive cellular remodeling that characterizes terminal differentiation. We found a robust alternative splicing program

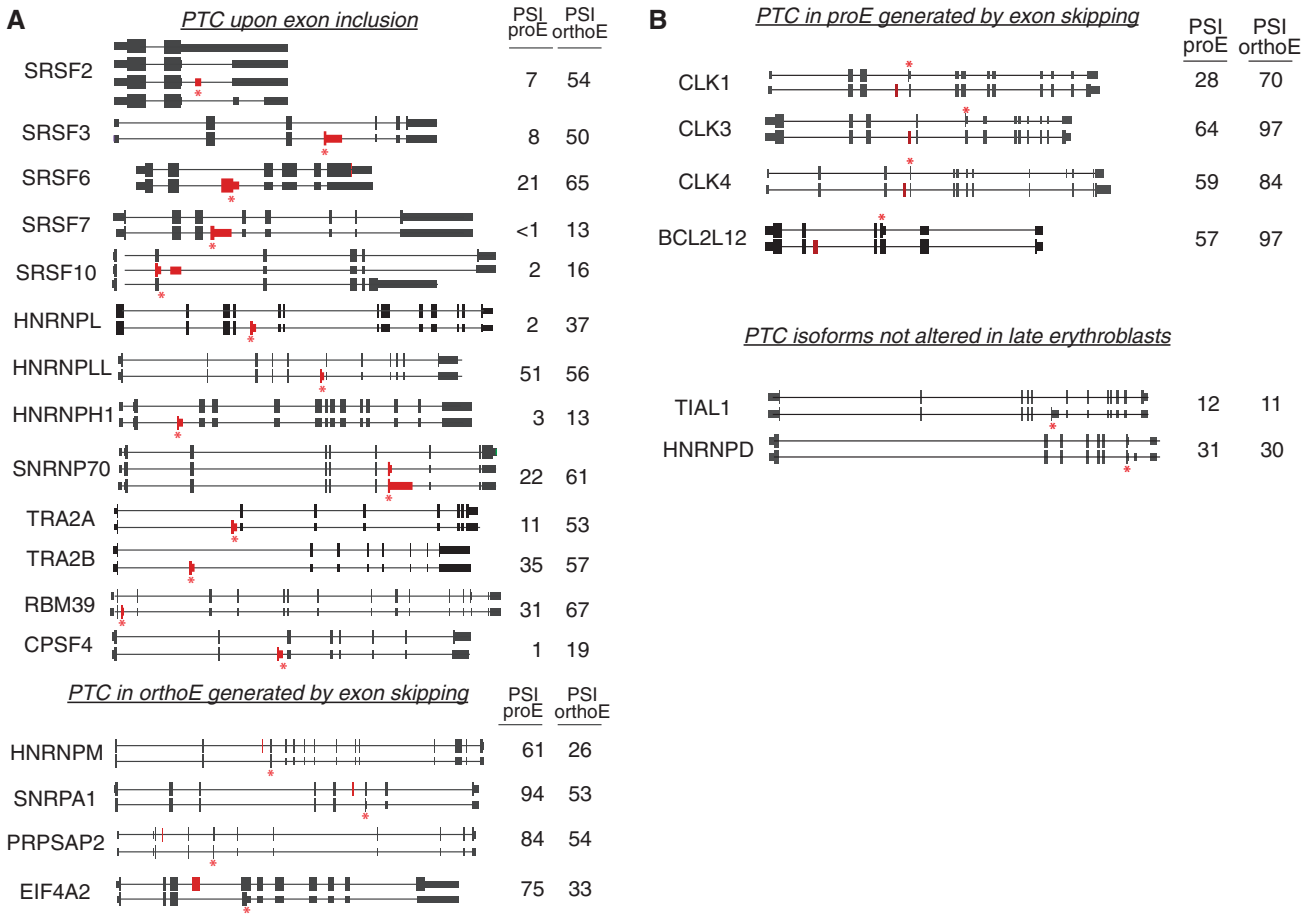

Figure 6. Differential expression of full-length and PTC transcript isoforms in differentiating erythroblasts. (**A**) RNA binding protein genes that increase expression of PTC isoforms in orthoE by exon inclusion (above) or by exon skipping (below). Exons in red contain stop codons (indicated by asterisks). PSI values for the PTC exons in proE and orthoE are shown at the right. (**B**) Genes in which PTC isoforms are more abundant in proE (above) or are not regulated in late erythroblasts (below).

enriched in genes that function in cell cycle regulation, organelle organization, chromatin structure and function and RNA processing, all of which undergo tremendous changes before enucleation. Most importantly, dynamic switches in alternative exon usage remodel the late erythroid transcriptome so as to generate new mRNA isoforms or to alter proportions of existing isoforms. These data support the hypothesis that a highly structured alternative splicing program is executed during terminal erythroid differentiation to regulate key biological processes as the cells prepare to enucleate and form mature red cells.

Previously, proper alternative splicing of protein 4.1 (EPB41) pre-mRNA was shown to be essential for generating a red cell membrane with the specialized mechanical properties required for survival in the circulation (48). The finding of abundant alternative splicing events in erythroblasts suggests many additional contributions to erythroid-specific functions. For example, regulators of alpha- and beta-globin gene expression such as ATRX and its interacting partners MECP2 and macroH2A1, as well as BCL11A and SOX6 all express multiple transcript isoforms whose importance has not been sufficiently evaluated. The GO categories enriched in alternative splicing events suggest additional roles in regulation of chromatin condensation, autophagy and enucleation in late erythroblasts, all key and unique events during late stages of terminal erythroid differentiation. We anticipate that some of these alternative splicing switches might be shared by other cell types undergoing the transition from proliferation to terminal differentiation; e.g. differences in relative abundance of H2AFY isoforms in tissues rich in proliferating versus differentiating cells (37) mirror the changes observed in erythroblasts. The erythroid differentiation system may be ideal for studying such splicing events.

Our studies suggest that a second major function for the erythroid alternative splicing program is post-transcriptional downregulation of gene expression in erythroblasts. We hypothesize that NMD-associated splicing switches, similar to switches in coding exons, are driven by changes in expression of splicing regulatory proteins during terminal erythroid differentiation that can up- or downregulate expression of PTC exons, according to physiological requirements for the encoded proteins. The marked changes in steady-state levels of PTC-containing transcripts during differentiation indicate that the mechanism regulating their production must be distinct from homeostatic (often autoregulatory) control mechanisms

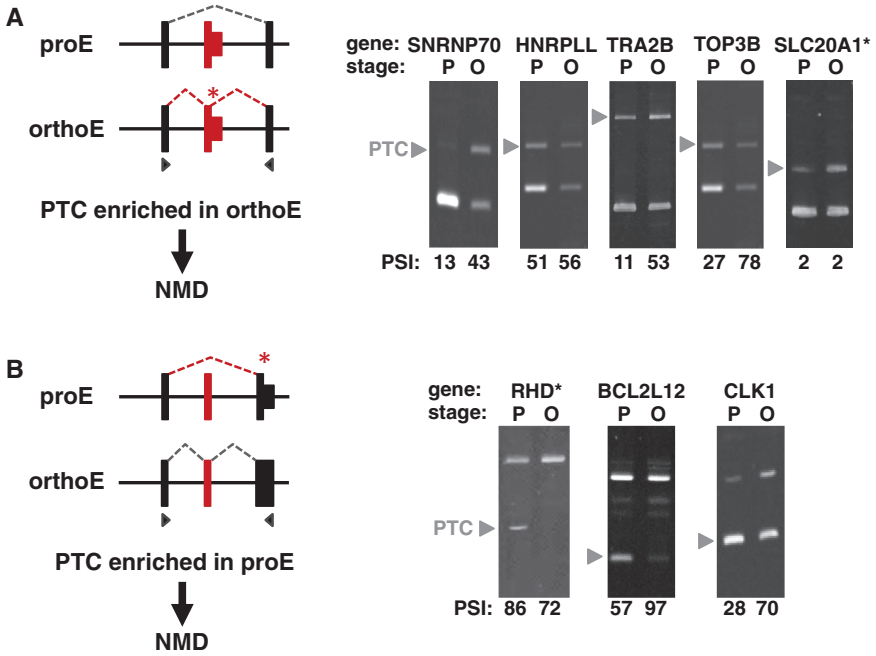

Figure 7. Validation of stage-specific PTC splicing events in human erythroblasts. (**A**) Transcripts in which PTC isoforms are more abundant in late erythroblasts due to upregulation of PTC-exons; (**B**) transcripts in which PTC transcripts are relatively more abundant in early erythroblasts due to exon skipping events that alter translational reading frame. Gene names and differentiation stage are indicated above each lane, while calculated PSI values are shown below each lane. Arrowheads indicate PCR bands representing PTC isoforms. * indicates genes for which PCR results indicate larger splicing changes than were predicted bioinformatically.

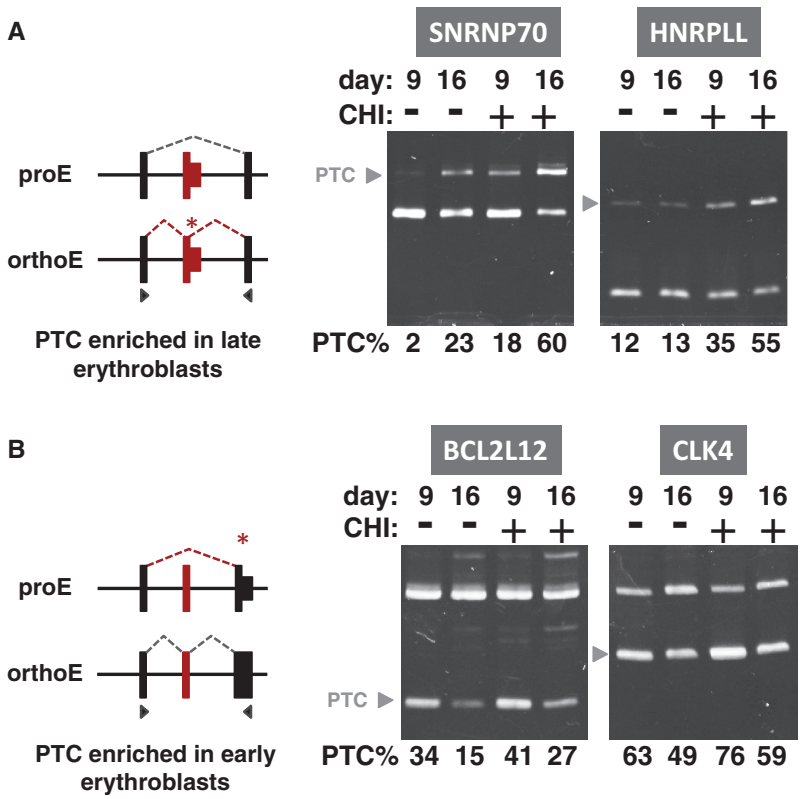

Figure 8. Effects of NMD inhibition on expression of PTC splicing events. (**A**) PTC transcripts produced by exon inclusion events were assayed in proE-enriched day 9 and polyE/orthoE-enriched day 16 erythroblast cultures without (−) or with (+) cycloheximide treatment. Cycloheximide dramatically increased the expression of PTC isoforms, especially at day 16, as expected if their abundance in untreated cells was limited by active NMD machinery. For SNRNP70, PCR detects the smaller of the two PTC exons depicted in Figure 7A, B. Cycloheximide effects on PTC-upon-exon skipping events. As in (A), cycloheximide increased expression of PTC isoforms, as expected, if their abundance in untreated cells was limited by NMD. The higher proportion of full-length transcripts predicted in late erythroblasts was maintained ±NMD inhibition. Gene names and culture stage are indicated above each lane, while PSI values shown below each lane were measured by densitometric analysis of stained gels. Arrowheads indicate PCR bands representing PTC isoforms.

previously ascribed to AS-NMD [(49) and references therein]. By maintaining active machinery for both alternative splicing and NMD, erythroblasts could coordinate these processes to limit expression in proliferating proE of differentiation functions required for remodeling in orthoE, while also downregulating in orthoE proliferative functions not needed in late erythroblasts. Differentiation-associated switches in PTC exons have already been proposed to operate in differentiating neurons, as, for example, changes in expression of splicing factor PTB can regulate switches in both PTC and non-PTC splicing events (50,51). The full impact of such pathways on erythroid biology may extend well beyond RNA binding proteins, as a substantial fraction of alternative splicing events affecting histone modifying enzymes (~42%, Supplementary Table S1) and many DNA binding proteins (results not shown) were predicted to induce NMD. Moreover, the current study likely underestimated or entirely missed PTC splicing switches that generate rapidly degraded transcripts, or splicing switches that resulted in retained introns; the latter occurs during late stages of granulocyte differentiation where they can suppress expression of selected genes (47). In fact, inspection of RNA reads mapping to the SF3B1 gene revealed increasing abundance of an intron retention event in late-stage erythroblasts (results not shown) that predicts reduced expression of this critical splicing factor implicated in myelodysplasia (4,52). Future studies will explore these potential expansions of AS-NMD by RNA-seq analysis of NMD-inhibited erythroblasts.

In conclusion, coordinated temporal control of alternative splicing plays a critical role in regulating structure and function of the erythroid transcriptome, likely having a major impact on orderly differentiation of erythroblasts. Analogous cell type-specific or differentiation-associated alternative splicing programs that are critical for cell identity have been identified in non-erythroid cell types including neurons, muscle cells, epithelial cells and T lymphocytes (3,53–55). Conversely, abnormal splicing factor expression as a consequence of human genetic disease or as a result of genetic knockdown in model organisms can profoundly alter splicing networks and have adverse phenotypic effects. Disturbance of programmed splicing networks underlies many complex human genetic diseases including myotonic dystrophy, the autoimmune neurologic disease paraneoplastic opsoclonus-myoclonus ataxia, facioscapulohumeral muscular dystrophy and amyotrophic lateral sclerosis (53,56–59). Moreover, a number of hematological cancers including myelodysplasia can be caused by mutations in splicing factor genes (4,52,60,61), and it has been shown that erythroblasts from myelodysplasia patients exhibit disordered erythropoiesis (62). Understanding the erythroblast splicing program will provide new insights into normal differentiation mechanisms and also into disordered differentiation in disease states.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

J.G.C., L.P., N.M. and X.A. designed the research; H.P., M.P., S.L.G., D.G. and J.L. performed research and analyzed data; and J.G.C., N.M., X.A., L.P. and H.P. wrote the article.

FUNDING

The National Institutes of Health (NIH) [DK094699 and DK032094]. Director, Office of Science, and Office of Biological & Environmental Research of the US Department of Energy under Contract No. DE-AC02-05CH1123. Funding for open access charge: NIH [DK094699].

Conflict of interest statement. None declared.

REFERENCES

1. Licatalosi,D.D. and Darnell,R.B. (2010) RNA processing and its regulation: global insights into biological networks. *Nat. Rev. Genet.*, **11**, 75–87.
2. Singh,R.K. and Cooper,T.A. (2012) Pre-mRNA splicing in disease and therapeutics. *Trends Mol. Med.*, **18**, 472–482.
3. Martinez,N.M., Pan,Q., Cole,B.S., Yarosh,C.A., Babcock,G.A., Heyd,F., Zhu,W., Ajith,S., Blencowe,B.J. and Lynch,K.W. (2012) Alternative splicing networks regulated by signaling in human T cells. *RNA*, **18**, 1029–1040.
4. Yoshida,K., Sanada,M., Shiraishi,Y., Nowak,D., Nagata,Y., Yamamoto,R., Sato,Y., Sato-Otsubo,A., Kon,A., Nagasaki,M. *et al.* (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, **478**, 64–69.
5. Cazzola,M., Rossi,M. and Malcovati,L. (2013) Biologic and clinical significance of somatic mutations of SF3B1 in myeloid and lymphoid neoplasms. *Blood*, **121**, 260–269.
6. Wan,Y. and Wu,C.J. (2013) SF3B1 mutations in chronic lymphocytic leukemia. *Blood*, **121**, 4627–4634.
7. Yamamoto,M.L., Clark,T.A., Gee,S.L., Kang,J.A., Schweitzer,A.C., Wickrema,A. and Conboy,J.G. (2009) Alternative pre-mRNA splicing switches modulate gene expression in late erythropoiesis. *Blood*, **113**, 3363–3370.
8. Discher,D., Parra,M., Conboy,J.G. and Mohandas,N. (1993) Mechanochemistry of the alternatively spliced spectrin-actin binding domain in membrane skeletal protein 4.1. *J. Biol. Chem.*, **268**, 7186–7195.
9. Horne,W.C., Huang,S.C., Becker,P.S., Tang,T.K. and Benz,E.J. Jr (1993) Tissue-specific alternative splicing of protein 4.1 inserts an exon necessary for formation of the ternary complex with erythrocyte spectrin and F-actin. *Blood*, **82**, 2558–2563.
10. Chasis,J.A., Coulombel,L., Conboy,J., McGee,S., Andrews,K., Kan,Y.W. and Mohandas,N. (1993) Differentiation-associated switches in protein 4.1 expression. Synthesis of multiple structural isoforms during normal human erythropoiesis. *J. Clin. Invest.*, **91**, 329–338.
11. Tang,T.K., Leto,T.L., Correas,I., Alonso,M.A., Marchesi,V.T. and Benz,E.J. Jr (1988) Selective expression of an erythroid-specific isoform of protein 4.1. *Proc. Natl Acad. Sci. USA*, **85**, 3713–3717.
12. Hou,V.C., Lersch,R., Gee,S.L., Ponthier,J.L., Lo,A.J., Wu,M., Turck,C.W., Koury,M., Krainer,A.R., Mayeda,A. *et al.* (2002) Decrease in hnRNP A/B expression during erythropoiesis mediates a pre-mRNA splicing switch. *EMBO J.*, **21**, 6195–6204.
13. Ponthier,J.L., Schluepen,C., Chen,W., Lersch,R.A., Gee,S.L., Hou,V.C., Lo,A.J., Short,S.A., Chasis,J.A., Winkelmann,J.C. *et al.* (2006) Fox-2 splicing factor binds to a conserved intron motif to promote inclusion of protein 4.1R alternative exon 16. *J. Biol. Chem.*, **281**, 12468–12474.
14. Yang,G., Huang,S.C., Wu,J.Y. and Benz,E.J. Jr (2005) An erythroid differentiation-specific splicing switch in protein 4.1R

mediated by the interaction of SF2/ASF with an exonic splicing enhancer. *Blood*, **105**, 2146–2153.
15. Yang,G., Huang,S.C., Wu,J.Y. and Benz,E.J. Jr (2008) Regulated Fox-2 isoform expression mediates protein 4.1R splicing during erythroid differentiation. *Blood*, **111**, 392–401.
16. Conboy,J.G., Shitamoto,R., Parra,M., Winardi,R., Kabra,A., Smith,J. and Mohandas,N. (1991) Hereditary elliptocytosis due to both qualitative and quantitative defects in membrane skeletal protein 4.1. *Blood*, **78**, 2438–2443.
17. Marchesi,S.L., Conboy,J., Agre,P., Letsinger,J.T., Marchesi,V.T., Speicher,D.W. and Mohandas,N. (1990) Molecular analysis of insertion/deletion mutations in protein 4.1 in elliptocytosis. I. Biochemical identification of rearrangements in the spectrin/actin binding domain and functional characterizations. *J. Clin. Invest.*, **86**, 516–523.
18. Hu,J., Liu,J., Xue,F., Halverson,G., Reid,M., Guo,A., Chen,L., Raza,A., Galili,N., Jaffray,J. *et al.* (2013) Isolation and functional characterization of human erythroblasts at distinct stages: implications for understanding of normal and disordered erythropoiesis *in vivo*. *Blood*, **121**, 3246–3253.
19. Flicek,P., Ahmed,I., Amode,M.R., Barrell,D., Beal,K., Brent,S., Carvalho-Silva,D., Clapham,P., Coates,G., Fairley,S. *et al.* (2013) Ensembl 2013. *Nucleic Acids Res.*, **41**, D48–D55.
20. Church,D.M., Schneider,V.A., Graves,T., Auger,K., Cunningham,F., Bouk,N., Chen,H.C., Agarwala,R., McLaren,W.M., Ritchie,G.R. *et al.* (2011) Modernizing reference genome assemblies. *PLoS Biol.*, **9**, e1001091.
21. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory–efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
22. Roberts,A. and Pachter,L. (2013) Streaming fragment assignment for real-time analysis of sequencing experiments. *Nat. Methods*, **10**, 71–73.
23. Mortazavi,A., Williams,B.A., McCue,K., Schaeffer,L. and Wold,B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods*, **5**, 621–628.
24. Trapnell,C., Williams,B.A., Pertea,G., Mortazavi,A., Kwan,G., van Baren,M.J., Salzberg,S.L., Wold,B.J. and Pachter,L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.*, **28**, 511–515.
25. Tusher,V.G., Tibshirani,R. and Chu,G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA*, **98**, 5116–5121.
26. Eden,E., Navon,R., Steinfeld,I., Lipson,D. and Yakhini,Z. (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*, **10**, 48.
27. Xu,J., Sankaran,V.G., Ni,M., Menne,T.F., Puram,R.V., Kim,W. and Orkin,S.H. (2010) Transcriptional silencing of {gamma}-globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev.*, **24**, 783–798.
28. Yi,Z., Cohen-Barak,O., Hagiwara,N., Kingsley,P.D., Fuchs,D.A., Erickson,D.T., Epner,E.M., Palis,J. and Brilliant,M.H. (2006) Sox6 directly silences epsilon globin expression in definitive erythropoiesis. *PLoS Genet.*, **2**, e14.
29. Clynes,D., Higgs,D.R. and Gibbons,R.J. (2013) The chromatin remodeller ATRX: a repeat offender in human disease. *Trends Biochem. Sci.*, **38**, 461–466.
30. Allis,C.D., Berger,S.L., Cote,J., Dent,S., Jenuwien,T., Kouzarides,T., Pillus,L., Reinberg,D., Shi,Y., Shiekhattar,R. *et al.* (2007) New nomenclature for chromatin-modifying enzymes. *Cell*, **131**, 633–636.
31. McKenney,R.J., Weil,S.J., Scherer,J. and Vallee,R.B. (2011) Mutually exclusive cytoplasmic dynein regulation by NudE-Lis1 and dynactin. *J. Biol. Chem.*, **286**, 39615–39622.
32. Keerthivasan,G., Liu,H., Gump,J.M., Dowdy,S.F., Wickrema,A. and Crispino,J.D. (2012) A novel role for survivin in erythroblast enucleation. *Haematologica*, **97**, 1471–1479.
33. Miura,K., Fujibuchi,W. and Unno,M. (2012) Splice variants in apoptotic pathway. *Exp. Oncol.*, **34**, 212–217.
34. Talbert,P.B. and Henikoff,S. (2010) Histone variants—ancient wrap artists of the epigenome. *Nat. Rev. Mol. Cell Biol.*, **11**, 264–275.
35. Gamble,M.J., Frizzell,K.M., Yang,C., Krishnakumar,R. and Kraus,W.L. (2010) The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. *Genes Dev.*, **24**, 21–32.
36. Buschbeck,M., Uribesalgo,I., Wibowo,I., Rue,P., Martin,D., Gutierrez,A., Morey,L., Guigo,R., Lopez-Schier,H. and Di Croce,L. (2009) The histone variant macroH2A is an epigenetic regulator of key developmental genes. *Nat. Struct. Mol. Biol.*, **16**, 1074–1079.
37. Pehrson,J.R., Costanzi,C. and Dharia,C. (1997) Developmental and tissue expression patterns of histone macroH2A1 subtypes. *J. Cell Biochem.*, **65**, 107–113.
38. Chen,Y., Chen,C.F., Chiang,H.C., Pena,M., Polci,R., Wei,R.L., Edwards,R.A., Hansel,D.E., Chen,P.L. and Riley,D.J. (2011) Mutation of NIMA-related kinase 1 (NEK1) leads to chromosome instability. *Mol Cancer*, **10**, 5.
39. White,M.C. and Quarmby,L.M. (2008) The NIMA-family kinase, Nek1 affects the stability of centrosomes and ciliogenesis. *BMC Cell Biol.*, **9**, 29.
40. Patil,M., Pabla,N., Ding,H.F. and Dong,Z. (2013) Nek1 interacts with Ku80 to assist chromatin loading of replication factors and S-phase progression. *Cell Cycle*, **12**, 2608–2616.
41. Yamahashi,Y., Saito,Y., Murata-Kamiya,N. and Hatakeyama,M. (2011) Polarity-regulating kinase partitioning-defective 1b (PAR1b) phosphorylates guanine nucleotide exchange factor H1 (GEF-H1) to regulate RhoA-dependent actin cytoskeletal reorganization. *J. Biol. Chem.*, **286**, 44576–44584.
42. Klebig,M.L., Wall,M.D., Potter,M.D., Rowe,E.L., Carpenter,D.A. and Rinchik,E.M. (2003) Mutations in the clathrin-assembly gene Picalm are responsible for the hematopoietic and iron metabolism abnormalities in fit1 mice. *Proc. Natl Acad. Sci. USA*, **100**, 8360–8365.
43. Scotland,P.B., Heath,J.L., Conway,A.E., Porter,N.B., Armstrong,M.B., Walker,J.A., Klebig,M.L., Lavau,C.P. and Wechsler,D.S. (2012) The PICALM protein plays a key role in iron homeostasis and cell proliferation. *PLoS One*, **7**, e44252.
44. Lareau,L.F., Inada,M., Green,R.E., Wengrod,J.C. and Brenner,S.E. (2007) Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature*, **446**, 926–929.
45. Saltzman,A.L., Kim,Y.K., Pan,Q., Fagnani,M.M., Maquat,L.E. and Blencowe,B.J. (2008) Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay. *Mol. Cell Biol.*, **28**, 4320–4330.
46. Ni,J.Z., Grate,L., Donohue,J.P., Preston,C., Nobida,N., O'Brien,G., Shiue,L., Clark,T.A., Blume,J.E. and Ares,M. Jr. (2007) Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev.*, **21**, 708–718.
47. Wong,J.J., Ritchie,W., Ebner,O.A., Selbach,M., Wong,J.W., Huang,Y., Gao,D., Pinello,N., Gonzalez,M., Baidya,K. *et al.* (2013) Orchestrated intron retention regulates normal granulocyte differentiation. *Cell*, **154**, 583–595.
48. Discher,D.E., Winardi,R., Schischmanoff,P.O., Parra,M., Conboy,J.G. and Mohandas,N. (1995) Mechanochemistry of protein 4.1's spectrin-actin-binding domain: ternary complex interactions, membrane binding, network integration, structural strengthening. *J. Cell Biol.*, **130**, 897–907.
49. Kalsotra,A. and Cooper,T.A. (2011) Functional consequences of developmentally regulated alternative splicing. *Nat. Rev. Genet.*, **12**, 715–729.
50. Licatalosi,D.D., Yano,M., Fak,J.J., Mele,A., Grabinski,S.E., Zhang,C. and Darnell,R.B. (2012) Ptbp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. *Genes Dev.*, **26**, 1626–1642.
51. Zheng,S., Gray,E.E., Chawla,G., Porse,B.T., O'Dell,T.J. and Black,D.L. (2012) PSD-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. *Nat. Neurosci.*, **15**, 381–388.
52. Papaemmanuil,E., Cazzola,M., Boultwood,J., Malcovati,L., Vyas,P., Bowen,D., Pellagatti,A., Wainscoat,J.S., Hellstrom-Lindberg,E., Gambacorti-Passerini,C. *et al.* (2011) Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N. Engl. J. Med.*, **365**, 1384–1395.

53. Ule,J., Jensen,K.B., Ruggiu,M., Mele,A., Ule,A. and Darnell,R.B. (2003) CLIP identifies Nova-regulated RNA networks in the brain. *Science*, **302**, 1212–1215.
54. Bland,C.S., Wang,E.T., Vu,A., David,M.P., Castle,J.C., Johnson,J.M., Burge,C.B. and Cooper,T.A. (2010) Global regulation of alternative splicing during myogenic differentiation. *Nucleic Acids Res.*, **38**, 7651–7664.
55. Warzecha,C.C., Jiang,P., Amirikian,K., Dittmar,K.A., Lu,H., Shen,S., Guo,W., Xing,Y. and Carstens,R.P. (2010) An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition. *EMBO J.*, **29**, 3286–3300.
56. Kanadia,R.N., Johnstone,K.A., Mankodi,A., Lungu,C., Thornton,C.A., Esson,D., Timmers,A.M., Hauswirth,W.W. and Swanson,M.S. (2003) A muscleblind knockout model for myotonic dystrophy. *Science*, **302**, 1978–1980.
57. Orengo,J.P., Chambon,P., Metzger,D., Mosier,D.R., Snipes,G.J. and Cooper,T.A. (2008) Expanded CTG repeats within the DMPK 3' UTR causes severe skeletal muscle wasting in an inducible mouse model for myotonic dystrophy. *Proc. Natl Acad. Sci. USA*, **105**, 2646–2651.
58. Pistoni,M., Shiue,L., Cline,M.S., Bortolanza,S., Neguembor,M.V., Xynos,A., Ares,M. Jr and Gabellini,D. (2013) Rbfox1 Downregulation and Altered Calpain 3 splicing by FRG1 in a mouse model of Facioscapulohumeral Muscular Dystrophy (FSHD). *PLoS Genet.*, **9**, e1003186.
59. Polymenidou,M., Lagier-Tourenne,C., Hutt,K.R., Bennett,C.F., Cleveland,D.W. and Yeo,G.W. (2012) Misregulated RNA processing in amyotrophic lateral sclerosis. *Brain Res.*, **1462**, 3–15.
60. Je,E.M., Yoo,N.J., Kim,Y.J., Kim,M.S. and Lee,S.H. (2013) Mutational analysis of splicing machinery genes SF3B1, U2AF1 and SRSF2 in myelodysplasia and other common tumors. *Int. J. Cancer*, **133**, 260–265.
61. Zhang,S.J., Rampal,R., Manshouri,T., Patel,J., Mensah,N., Kayserian,A., Hricik,T., Heguy,A., Hedvat,C., Gonen,M. *et al.* (2012) Genetic analysis of patients with leukemic transformation of myeloproliferative neoplasms shows recurrent SRSF2 mutations that are associated with adverse outcome. *Blood*, **119**, 4480–4485.
62. Liu,J., Zhang,J., Ginzburg,Y., Li,H., Xue,F., De Franceschi,L., Chasis,J.A., Mohandas,N. and An,X. (2013) Quantitative analysis of murine terminal erythroid differentiation *in vivo*: novel method to study normal and disordered erythropoiesis. *Blood*, **121**, e43–49.