Proteomic/transcriptomic analysis of erythropoiesis

Marjorie Brand[a,b] *and Jeffrey A. Ranish*[c]

Purpose of review
Erythropoiesis is a hierarchical process by which hematopoietic stem cells give rise to red blood cells through gradual cell fate restriction and maturation. Deciphering this process requires the establishment of dynamic gene regulatory networks (GRNs) that predict the response of hematopoietic cells to signals from the environment. Although GRNs have historically been derived from transcriptomic data, recent proteomic studies have revealed a major role for posttranscriptional mechanisms in regulating gene expression during erythropoiesis. These new findings highlight the need to integrate proteomic data into GRNs for a refined understanding of erythropoiesis.

Recent findings
Here, we review recent proteomic studies that have furthered our understanding of erythropoiesis with a focus on quantitative mass spectrometry approaches to measure the abundance of transcription factors and cofactors during differentiation. Furthermore, we highlight challenges that remain in integrating transcriptomic, proteomic, and other omics data into a predictive model of erythropoiesis, and discuss the future prospect of single-cell proteomics.

Summary
Recent proteomic studies have considerably expanded our knowledge of erythropoiesis beyond the traditional transcriptomic-centric perspective. These findings have both opened up new avenues of research to increase our understanding of erythroid differentiation, while at the same time presenting new challenges in integrating multiple layers of information into a comprehensive gene regulatory model.

Keywords
absolute abundance, gene regulatory networks, proteomics, quantitative mass spectrometry, transcription factors

INTRODUCTION

Erythropoiesis is a gradual process leading to the formation of red blood cells from hematopoietic stem cells (HSCs) [1–3]. This process is driven by combinations of transcription factors (TFs) that establish lineage-specific gene expression programs while simultaneously inhibiting nonerythroid hematopoietic lineages [4,5]. The role of TFs in erythropoiesis can be best understood in the context of gene regulatory networks (GRNs) that integrate multiple types of information to describe and predict the response of hematopoietic cells to signals from the environment, leading to the production of red blood cells in sufficient quantities to respond to the demands of the organism [6▪]. Although invaluable to our understanding of erythropoiesis, GRNs have so far been derived almost exclusively from transcriptomic data [7,8] that have limited their usefulness because proteins (not transcripts) are the drivers of transcriptional regulatory mechanisms. Importantly, there is only a limited correlation between the abundances of mRNA and their corresponding proteins [9–15, 16▪,17▪,18]. Furthermore, we recently showed that master regulators of erythropoiesis (e.g. GATA1, KLF1, KLF3, GFI1B, TAL1) display large discrepancies in their protein versus mRNA dynamics [17▪]. Several, nonmutually exclusive posttranscriptional mechanisms can explain these discrepancies. For example, ribosome profiling experiments showed that GATA1 protein production is regulated for a large part at the translational level [15] whereby a short and unstructured 5′UTR transcript renders *GATA1* translation particularly sensitive to changes in ribosome levels [19,20▪]. Translational efficiency in erythropoiesis

[a]Sprott Center for Stem Cell Research, Ottawa Hospital Research Institute, [b]Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada and [c]Institute for Systems Biology, Seattle, Washington, USA

Correspondence to Marjorie Brand, Sprott Center for Stem Cell Research, Ottawa Hospital Research Institute, Ottawa, Ontario K1H8L6, Canada. Tel: +1 613 737 7700; e-mail: mbrand@ohri.ca and Jeffrey Ranish, Institute for Systems Biology, Seattle, WA 98109, USA. Tel: +1 206 732 1357; e-mail: jeff.ranish@isbscience.org

Curr Opin Hematol 2021, 28:150–157
DOI:10.1097/MOH.0000000000000647

KEY POINTS

- Major discrepancies between mRNA and protein abundances for master regulators of erythropoiesis.
- Absolute quantification of proteins by mass spectrometry uniquely provides information about the stoichiometry and dynamics of native proteins in erythropoiesis.
- Direct quantitative protein measurements (copy number per cell) provide key information for building quantitative models of gene regulatory networks in erythropoiesis.
- The nucleus is a highly repressive environment with corepressors that are two orders of magnitude more abundant than coactivators, likely a general principle of transcriptional regulation during cellular differentiation.

can also be modulated by RNA binding proteins such as RBM38 that promotes translation of select erythroid transcripts [21] or LIN28B that inhibits translation of the gamma-globin repressor BCL11A [22]. Furthermore, protein levels can be directly controlled through interaction with other proteins as recently shown for KLF1 which is protected from proteasomal degradation by the G protein pathway suppressor 2 (GPS2) protein [23▪]. Although multiple posttranscriptional mechanisms are likely to participate in regulating protein levels, it has been shown that the rate of protein translation in erythroid progenitors is greatly elevated [24▪] suggesting that translation plays a particularly important role during erythropoiesis. Further supporting this possibility, mass spectrometry (MS)-based analyses have identified a global decrease in ribosomal proteins in Diamond–Blackfan anemia (DBA) [19], a genetic disease with mutations in ribosomal proteins genes whereby patients display a selective defect in erythropoiesis [25].

Taken together, the above studies underline a critical need for direct protein measurements to better understand the regulation of erythropoiesis. Here, we review recent proteomic studies that have furthered our understanding of erythropoiesis with a focus on quantitative MS approaches. We also provide a perspective on current challenges and exciting future opportunities in the field.

TEXT OF REVIEW

Proteomic studies in erythropoiesis: quantitative mass spectrometry approaches

MS-based proteomics is a powerful technique for studying erythropoiesis because it is relatively unbiased, and thousands of proteins can be identified in a typical analysis of samples derived from whole cells or subcellular compartments. Important for studying dynamic processes such as erythropoiesis, MS can also measure quantitative changes in protein levels using either relative or absolute quantification approaches. In the following paragraphs, we highlight studies that have used relative or absolute MS-based quantification strategies to provide important insights in our understanding of erythropoiesis.

Relative quantification approaches

In a relative quantification experiment, changes in the abundances of proteins across samples (for example at different time-points during erythropoiesis, or between healthy and disease cells) are determined. Relative quantification can be achieved by label-free or isotopic labeling based-approaches. In the label-free approach, each sample is analyzed separately by MS and the relative levels of the proteins in the samples are determined by comparing either the peptide ion signals or the number of identified peptide spectra corresponding to each protein in the different samples. Quantification can also be performed through differential isotopic labeling of proteins or peptides that are then combined prior to MS analysis during which relative protein abundances are determined by comparing the ion signals of isotopically labeled sibling peptides (or their fragment ions). In one of the very first applications of the isotopic labeling quantitative MS approach in 2004, we used the isotope-coded affinity tag method to identify a switch in the proteins that interact with the TF MafK from corepressors to coactivators, an event critical for ß-globin gene activation during erythroid differentiation [26]. More recently, a number of laboratories have used the iTRAQ isobaric labeling technique for erythropoiesis studies. For example, the Xu lab identified mitochondrial factors that play critical roles in erythroid specification [15] as well as proteins that bind to the ß-globin locus [27]. Also using iTRAQ, the Mayeux lab compared the protein content of pyrenocytes and reticulocytes, revealing that proteasome particles largely segregate to the reticulocyte [14]. We have also used iTRAQ to obtain a global view of quantitative changes in nuclear protein levels at multiple time points along the erythroid lineage from hematopoietic stem/progenitor cells [17▪]. Finally, using a different isobaric labeling approach ('tandem mass tags' or TMT [28]), the Frayne lab identified differentially expressed proteins between adult and cord blood red blood cells [29], the Sankaran lab found a global reduction in ribosome proteins levels in erythroid progenitors

from patients with DBA [19] and the Blanc lab identified dexamethasone targets that regulate erythroid progenitors' proliferation [30]. Thus, proteomic studies based on relative quantification of proteins have played a major role in revealing molecular processes that are important for erythropoiesis in health and disease.

Absolute quantification approaches

Despite their usefulness in highlighting important aspects of protein dynamics, MS approaches that rely on relative protein quantification do not provide information on protein copy number and thus have several limitations. First, they do not inform us on protein availability in a specific sample (e.g. we do not know how many molecules of TFs are present in a cell compared to the number of DNA binding sites). Furthermore, they do not measure protein stoichiometry, information that is critical for understanding protein complex composition and formation, or competition between TFs that drive cell fate decisions in erythropoiesis. Finally, relative quantification has limited utility for quantitative modeling of GRNs. Absolute measurements, on the other hand, permit the determination of the number of molecules of each protein in a sample and have therefore an extended range of applications (see examples below).

Absolute quantification of proteins by MS can be achieved by two main approaches: (1) by using isotopically labeled peptide standards or (2) by label-free methods (for review, [13]). The first approach, stable isotope dilution (SID), employs stable isotope labeled (SIL) peptides corresponding to peptides derived from proteins of interest that are spiked into the biological sample at known concentrations. Thus, the absolute concentration of the proteins of interest is determined by comparing the ion signals produced by the SIL peptides to the signals produced by the sample peptides. This approach is considered the gold standard for protein quantification by MS because each peptide is quantified by using its own internal control (i.e. an isotopically labeled version of itself). Because the ion signal that each peptide produces in the mass spectrometer can be affected by factors such as amino acid sequence and sample composition, the SIL peptide approach is less prone to quantification inaccuracies due to these issues. On the other hand, because the SIL approach is typically based on a few peptides per protein, quantification can be affected by issues such as PTMs or differences in digestion efficiencies. This approach is commonly used in conjunction with targeted MS approaches, such as selected reaction monitoring (SRM), which are well suited for reproducible and quantitative measurements of a defined set of analytes over a wide dynamic range of abundances [31,32,33▪]. However due to the cost of synthetic SIL peptides, this approach is typically used for the quantification of relatively small numbers (10–50) of proteins. Alternative approaches for production of SIL peptides such as recombinant expression of artificial proteins composed of target peptides or full-length proteins can reduce the cost of SIL peptides [34–38]. In addition, chemical tagging approaches that permit sample multiplexing such TMT or iTRAQ, can be combined with spiked-in peptide/protein standards and high throughput MS approaches, to improve the throughput of the analysis [39,40].

Label-free approaches permit quantification of large numbers of proteins in a high-throughput manner. Since they do not require SIL peptides, label-free approaches are not as expensive as the SID approach, and they are less time-consuming. Label-free quantification is based on either the number of identified peptides corresponding to a protein or the ion signals that the peptides corresponding to a protein produce in the mass spectrometer [41]. To estimate the absolute abundance of each protein, the label-free values can be compared to standards of known abundance [11,41] or to the total protein amount of the sample. In this later method, called the total protein approach (TPA) [42], the copy number for each protein is calculated from the ratio of each protein's total peptide ion signal divided by the total ion signal of the measured proteome with knowledge of the proteins molecular weight, Avogadro's number and by the protein content of a single cell. In an extension of this approach, called the 'proteomic ruler,' the ion signals derived from the histone proteins are used as a proxy for the mass of DNA. This value allows estimation of cell number which in turn allows scaling the protein abundances by cell number. The authors found that the accuracy of TPA was best for proteins quantified with more than a few peptides. Overall, label-free methods provide a rapid and cost-effective approach to estimate absolute abundances for large numbers of proteins without the need for SIL peptides Care must be taken however for proteins quantified by only a few peptides, which typically corresponds to less abundant proteins due to the potential for different peptides to produce different MS signals per mole. Furthermore, potential changes in cellular histone content should be taken into consideration when using the proteomic ruler method [43,44].

Insights from absolute quantification of proteins using label-free mass spectrometry

Label-free MS approaches were used in several large-scale studies of erythropoiesis [14,16▪,45▪]. All these studies measured proteins from whole-cell lysates

(made of a mixture of cytoplasmic and nuclear proteins) that have been prepared from either purified or synchronized cell populations ranging from erythroid progenitors to terminally differentiated orthochromatic erythroblasts. Overall, these studies identified and quantified 6,000 to 8,000 total proteins at different stages of differentiation, including proteins from the nucleus, the cytoplasm, the plasma membrane (including solute carrier transporters of metal ions), mitochondria, and even secreted proteins. The data provide a previously unappreciated scaling of these various classes of factors. Overall, these global analyses revealed a gradual progression of protein clusters across the different stages of erythropoiesis and identified a decrease in the total amount of cellular proteins during terminal erythroid differentiation. Interestingly, comparative analyses between protein levels and transcripts revealed a global dampening of quantitative variations at the proteome level suggesting that much of transcriptional noise is buffered at the protein level. Finally, comparative analyses between the proteome of primary erythroid cells with that of cell lines commonly used for studying erythroid differentiation allowed the authors to pinpoint the exact stage of differentiation at which these cell lines are arrested, showing that global comparative analyses of proteins are useful to identify specific stages of differentiation [16▪]. Although the conclusions from these studies are mostly in agreement, we note that Gautier *et al.* did not detect a significant decrease in overall cellular histone contents at the end of erythroid differentiation [14,16▪] whereas Karayel *et al.* were able to detect this change [45▪] first suggested in previous studies [43,44]. This inconsistency remains to be addressed. Although these data provide valuable information to detect global changes in multiple classes of proteins during erythropoiesis, the limited coverage of low abundance proteins such as TFs makes this approach less appropriate for incorporation into GRNs.

Insights from absolute quantification of transcription-related proteins using selected reaction monitoring-based mass spectrometry with stable isotope labeled peptides

To identify quantitative changes in the stoichiometry of TFs and cofactors during human erythropoiesis, we extracted nuclear proteins at multiple stages of erythropoiesis from hematopoietic stem/progenitor cells and we used SRM-based targeted MS coupled with the spiking of known amounts of isotopically labeled peptides to quantify over one hundred proteins, including TFs, cofactors, chromatin remodeling enzymes and subunits of the general transcription machinery [17▪,33▪,46]. These analyses led to several major findings. First, we found that the correlation between protein and RNA levels is the lowest at the early stages of differentiation in hematopoietic stem/progenitor cells [17▪], consistent with the previous finding that these cells have a low protein translation rate [47]. Furthermore, this result is in agreement with other MS-based studies using purified hematopoietic progenitor populations of both human [18] and mouse [48] origin. Importantly, we were able to measure for the first time, changes in the stoichiometry between antagonistic TFs in hematopoietic progenitors, showing that the lineage specifying (LS) TFs KLF1 and FLI1 are both present at about 3,000 copies per cell in hematopoietic progenitors and that KLF1 protein gradually becomes more abundant as the cells progress along the erythroid lineage [46]. Interestingly, our data showed major discrepancies in the dynamics of transcripts versus proteins during the course of erythropoiesis, leading to reconsideration of some previously accepted models that were based on transcripts measurements. For example, whereas the switch from GATA2 to GATA1 genomic binding that occurs at the proEB stage has traditionally been thought to be driven by GATA1 becoming more abundant than GATA2, this view is not compatible with the proteomic data. Indeed, at the protein level, GATA1 is more abundant than GATA2 from the early stages of hematopoiesis. Thus, GATA2 is able to bind to its target genes despite an excess of GATA1, and the GATA2/GATA1 switch is likely to be mediated by a decrease in GATA2 protein rather than just a competition with GATA1 [17▪].

Perhaps the most important finding from our study was the unexpected large difference in abundances between transcriptional co-repressors and co-activators. Specifically, we observed that in the nucleus, co-repressors are on average 2 orders of magnitude more abundant than co-activators with TFs being present at intermediate levels (Fig. 1a). Interestingly, these differences exist at the protein, not the RNA level and can be at least partially explained by the instability of co-activator proteins [17▪]. This finding has major implications for our understanding of transcriptional regulation during erythropoiesis in general by implying that DNA binding TFs must compete with each other to recruit rare co-activators in the context of a highly repressive nuclear environment (Fig. 1b). Furthermore, a high co-repressor / co-activator ratio is likely to play an important role in maintaining oligopotency and regulating cell fate decisions in progenitors by preventing inappropriate expression of genes from mutually exclusive lineages (Fig. 2). Consistent with this, one of the most abundant factors we quantified

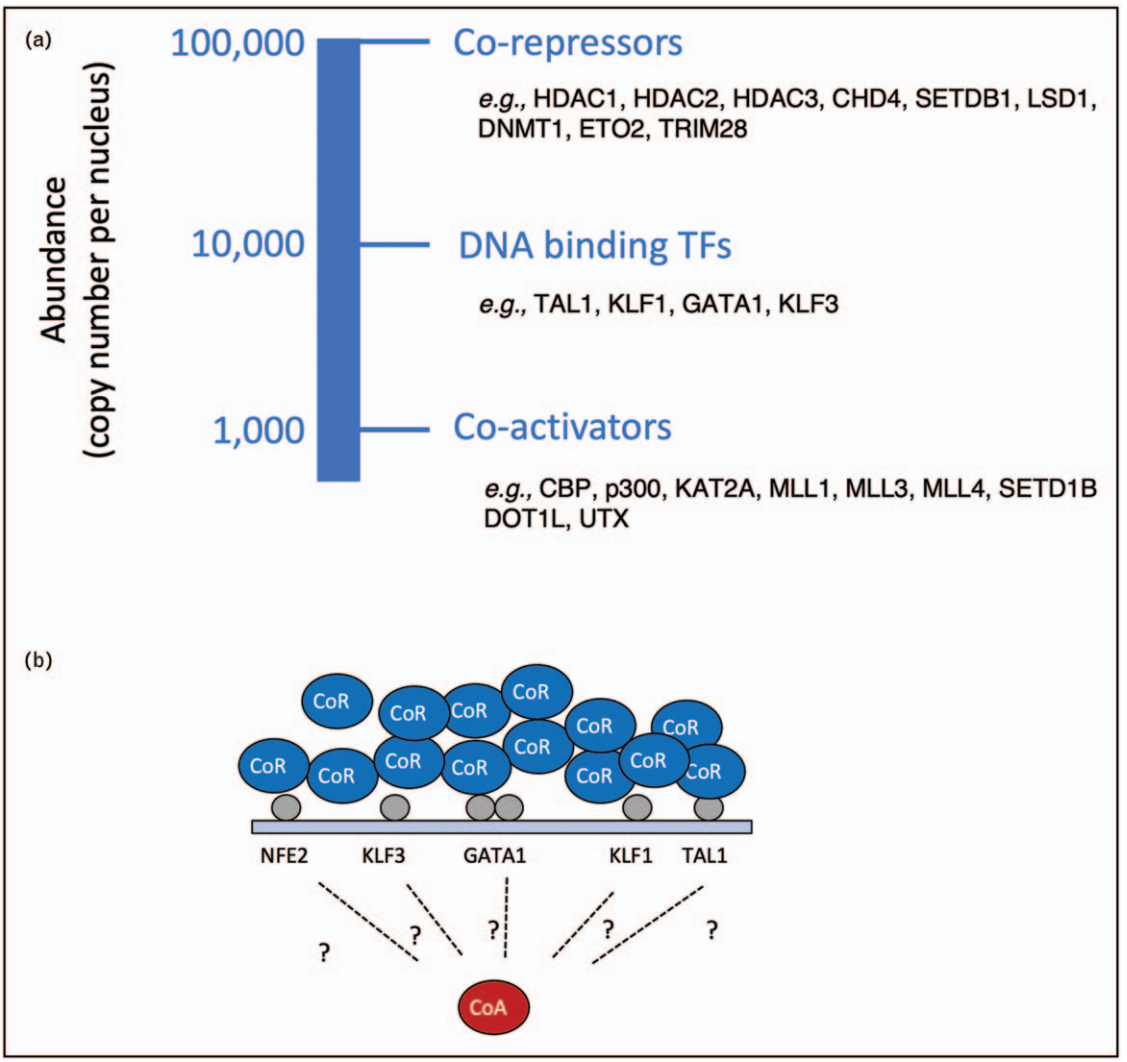

FIGURE 1. Stoichiometry of different classes of proteins in the nucleus. (a) Abundance of different classes of proteins in the nucleus, including co-activators, transcription factors and co-repressors. (b) The high abundance of co-repressors facilitates their recruitment to target genes whereas transcription factors must compete to recruit rare co-activators.

(i.e. CHD4, over 400,000 copies per nucleus) is a subunit of the NuRD co-repressor complex that plays a critical role in repressing the fetal γ-globin gene [49] and suppressing transcriptional noise [50]. It is likely that the repressive nuclear environment is a general principle of cellular differentiation.

Finally, it should be noted that even though co-activators are rare in the nucleus, their protein levels can vary dramatically during erythropoiesis [17▪] suggesting that they may have specific roles at different stages of differentiation. Consistent with this, it has been shown that histone-modifying enzymes, including the histone acetyltransferase MOF [51▪▪] or the histone H3K36 methyltransferase NSD1 [52] have erythroid-specific functions. Although our study provided quantitative information on a diverse set of representative subunits from transcriptional co-regulatory complexes, comprehensive quantification of complex subunits remains to be performed in a systematic manner to assess potential changes in stoichiometry during erythropoiesis.

Proteomic studies in erythropoiesis: future prospects and challenges

Despite the considerable amount and variety of proteomic data available, additional measurements are necessary to fully comprehend gene regulation in erythropoiesis. These measurements ought to be quantitative and aim to include all expressed proteins and their proteoforms, including posttranslational modifications (PTMs) [53▪]. Indeed, while a global phospho-proteome study has been published in erythroid cells [45▪], additional PTMs such as

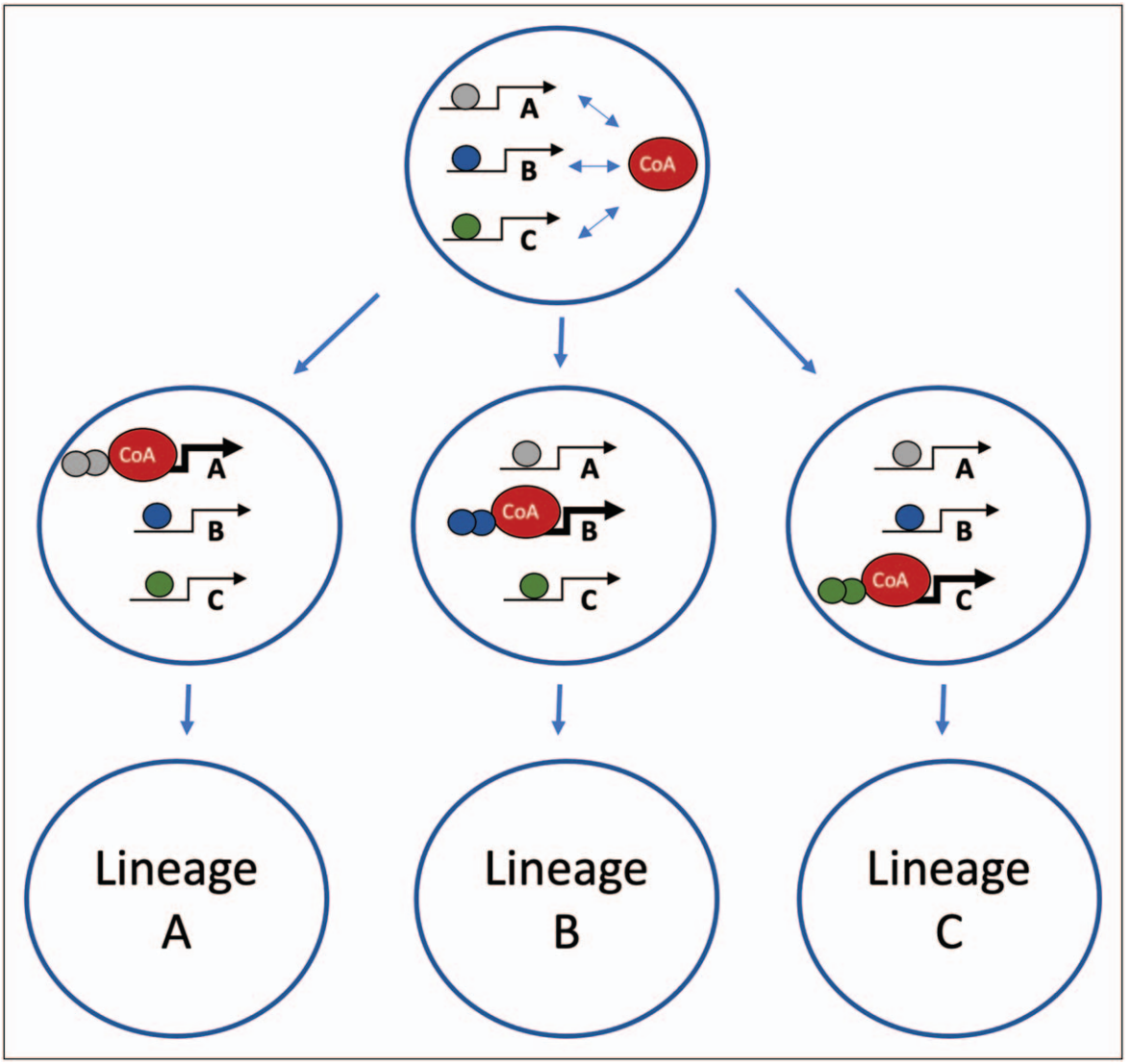

FIGURE 2. Proposed model of cell fate decision facilitated by the stoichiometry between transcription factors, co-repressors and co-activators. This model is based on the high abundance of co-repressors (not shown), and the low abundance of co-activators in the nucleus as determined by quantitative mass spectrometry. In the multipotent progenitor cell (top row), the low abundance of co-activators coupled to the large excess of co-repressors prevent high level activation of genes from mutually exclusive lineages A, B and C. As the progenitors differentiate (second row), the abundance of lineage-specifying transcription factors (colored circles) increases, which facilitates the recruitment of co-activators to lineage-specific genes at the expense of genes from other lineages.

acetylation, sumoylation and others remain to be systematically identified. Furthermore, a recent study highlighted the effect of trace metals such as zinc in modifying the proteome of erythroid cells [54], an interesting area to pursue. Given the complexity of the proteome (i.e. >100 M predicted proteoforms derived from 83,000 transcripts and 20,000 genes [53▪], and the interactions amongst these proteins), much remains to be done to reveal the full proteome and its changes during erythropoiesis.

Beside generating new proteomic data, two major challenges remain: (1) integrating these data within other data types such as DNA binding (e.g., by chromatin-immunoprecipitation) into a comprehensive and predictive GRN, and (2) moving MS-based proteomics into the single-cell arena.

Data integration and gene regulatory networks

Using quantitative protein measurements by SRM and ordinary differential equations, we recently built the first temporal GRN for erythroid commitment that integrates quantitative changes in protein and mRNA abundances of key TFs [17▪]. Notably,

our model accurately recapitulated a number of previously known regulatory links, including the cross-antagonisms between lineage specifying (LS)-TFs in their correct sequential order. Importantly the model revealed that these cross-antagonisms are quantitatively imbalanced and that these imbalances become more pronounced with time, further supporting the continuous and quantitative nature of cell fate decisions. Despite this initial success, much remains to be done in this area, including integration of information about co-activators and co-repressors, PTMs, protein-protein interactions as well as protein-DNA and protein-RNA interactions into a refined model of erythropoiesis. The numerous challenges that remain in data integration and quantitative modeling of GRNs have been reviewed recently [6▪,55].

Single-cell proteomics: are we there yet?

In the past 5–6 years, single-cell transcriptomics has revolutionized our understanding of hematopoiesis and erythropoiesis [5]. However, single-cell proteomic technology has not kept pace with transcriptomics, which is largely due to a lack of sensitivity of current protein-based technologies. And while antibody-based single cell approaches such as CyTOF have the sensitivity to permit measurements of low abundance proteins like TFs, they are limited to proteins for which suitable antibodies are available. Furthermore, antibody-based approaches do not provide information about absolute protein amounts unless calibrated with standards of known abundance [46].

Importantly, recent studies suggest that single-cell proteomics by MS (sc-MS) may soon be an achievable goal even for low abundance proteins. For example, recent advances in MS instrumentation, sample processing, and experimental design have allowed the identification of ~1000 protein groups from single cells using both label-free [56] and isobaric labeling approaches [57▪,58▪▪]. Furthermore, ~6000 protein groups have been identified in samples derived from a few hundred cells [59]. Notably, the Slavov laboratory recently demonstrated that a handful of TFs could by identified and quantified by sc-MS in macrophages [58▪▪]. Finally, temporal trajectories were recently established from sc-MS data of murine macrophages during lipopolysaccharide stimulation [60]. Based on these new exciting results, it is entirely conceivable that, with further advances in sample processing, experimental design, and computational approaches, MS-based proteomics will soon permit large-scale proteome measurements in single cells, opening new doors to understanding transcriptional regulation of erythropoiesis and other dynamic cellular processes.

CONCLUSION

Recent proteomic studies have considerably expanded our knowledge of erythropoiesis beyond the traditional transcriptomic-centric perspective. These findings have both opened up new avenues of research to increase our understanding of erythroid differentiation while presenting new challenges in integrating multiple layers of information into a comprehensive gene regulatory model.

Acknowledgements

None.

Financial support and sponsorship

The work was supported by NIH DK098449 (M.B. and J.A.R.), CIHR MOP89834, MOP343603 and CEEHRC Initiative (M.B.), Canadian Cancer Research Society (M.B.), NIH GM110064, GM136974, and HL133678 (J.A.R.).

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:
- ▪ of special interest
- ▪▪ of outstanding interest

1. Palis J. Primitive and definitive erythropoiesis in mammals. Front Physiol 2014; 5:3.
2. Bresnick EH, Hewitt KJ, Mehta C, *et al.* Mechanisms of erythrocyte development and regeneration: implications for regenerative medicine and beyond. Development 2018; 145:dev151423.
3. Chen L, Wang J, Liu J, *et al.* Dynamic changes in murine erythropoiesis from birth to adulthood: implications for the study of murine models of anemia. Blood Adv 2021; 5:16–25.
4. Graf T, Enver T. Forcing cells to change lineages. Nature 2009; 462:587–594.
5. Brand M, Morrissey E. Single-cell fate decisions of bipotential hematopoietic progenitors. Curr Opin Hematol 2020; 27:232–240.
6. ▪ Rothenberg EV. Causal gene regulatory network modeling and genomics: second-generation challenges. J Comput Biol 2019; 26:703–718.
This review provides a very complete summary of the challenges and opportunities in building predictive gene regulatory models.
7. Swiers G, Patient R, Loose M. Genetic regulatory networks programming hematopoietic stem cells and erythroid lineage specification. Dev Biol 2006; 294:525–540.
8. Novershtern N, Subramanian A, Lawton LN, *et al.* Densely interconnected transcriptional circuits control cell states in human hematopoiesis. Cell 2011; 144:296–309.
9. Tian Q, Stepaniants SB, Mao M, *et al.* Integrated genomic and proteomic analyses of gene expression in Mammalian cells. Mol Cell Proteom 2004; 3:960–969.
10. Vogel C, Abreu Rde S, Ko D, *et al.* Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. Mol Syst Biol 2010; 6:400.
11. Schwanhausser B, Busse D, Li N, *et al.* Global quantification of mammalian gene expression control. Nature 2011; 473:337–342.
12. Li JJ, Bickel PJ, Biggin MD. System wide analyses have underestimated protein abundances and the importance of transcription in mammals. PeerJ 2014; 2:e270.
13. Liu Y, Beyer A, Aebersold R. On the dependency of cellular protein levels on mRNA abundance. Cell 2016; 165:535–550.
14. Gautier EF, Ducamp S, Leduc M, *et al.* Comprehensive proteomic analysis of human erythropoiesis. Cell Rep 2016; 16:1470–1484.
15. Liu X, Zhang Y, Ni M, *et al.* Regulation of mitochondrial biogenesis in erythropoiesis by mTORC1-mediated protein translation. Nat Cell Biol 2017; 19:626–638.

16. Gautier EF, Leduc M, Ladli M, *et al.* Comprehensive proteomic analysis of ■ murine terminal erythroid differentiation. Blood Adv 2020; 4:1464–1477.
This study provides a useful resource of the proteome and transcriptome during murine terminal erythroid differentiation. Furthermore, through a comparative analysis of the proteomes of primary cells and cell lines, the authors showed that HUDEP cells are the closest relatives to primary erythroid cells.
17. Gillespie MA, Palii CG, Sanchez-Taltavull D, *et al.* Absolute quantification of ■ transcription factors reveals principles of gene regulation in erythropoiesis. Mol Cell 2020; 78:960–974e911.
Using absolute quantification of proteins during erythropoiesis, this study showed that there are large discrepancies between proteins and mRNA for master regulators of differentiation. The authors have also showed that transcriptional corepressors are two orders magnitude more abundant than coactivators. Futhermore, the study describes the first temporal regulatory network for erythroid specification that integrates dynamic changes in protein and transcript levels for master regulators.
18. Amon S, Meier-Abt F, Gillet LC, *et al.* Sensitive quantitative proteomics of human hematopoietic stem and progenitor cells by data-independent acquisition mass spectrometry. Mol Cell Proteom 2019; 18:1454–1467.
19. Khajuria RK, Munschauer M, Ulirsch JC, *et al.* Ribosome levels selectively regulate translation and lineage commitment in human hematopoiesis. Cell 2018; 173:90–103e119.
20. Boussaid I, Le Goff S, Floquet C, *et al.* Integrated analyses of translatome and ■ proteome identify the rules of translation selectivity in RPS14-deficient cells. Haematologica 2020; Early view Apr 23. https://doi.org/10.3324/haematol.2019.239970.
This study identified several characteristics explaining the translational selectivity of erythroid factors in the presence of reduced levels of ribosomes, including the 3'UTR structure, transcript length and codon usage.
21. Alvarez-Dominguez JR, Zhang X, Hu W. Widespread and dynamic translational control of red blood cell development. Blood 2017; 129:619–629.
22. Basak A, Munschauer M, Lareau CA, *et al.* Control of human hemoglobin switching by LIN28B-mediated regulation of BCL11A translation. Nat Genet 2020; 52:138–145.
23. Ma WB, Wang XH, Li CY, *et al.* GPS2 promotes erythroid differentiation by ■ control of the stability of EKLF protein. Blood 2020; 135:2302–2315.
This study shows that the G-protein pathway suppressor 2 serves as a chaperone for EKLF protein, increasing its stability, thereby controlling erythroid differentiation.
24. Magee JA, Signer RAJ. Developmental stage-specific changes in protein ■ synthesis differentially sensitize hematopoietic stem cells and erythroid progenitors to impaired ribosome biogenesis. Stem Cell Rep 2021; 16:20–28.
This study shows that adult erythroid progenitors display a higher protein translation rate than their fetal counterparts.
25. Iskander D, Psaila B, Gerrard G, *et al.* Elucidation of the EP defect in Diamond-Blackfan anemia by characterization and prospective isolation of human EPs. Blood 2015; 125:2553–2557.
26. Brand M, Ranish JA, Kummer NT, *et al.* Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. Nat Struct Mol Biol 2004; 11:73–80.
27. Liu X, Zhang Y, Chen Y, *et al.* CAPTURE: in situ analysis of chromatin composition of endogenous genomic loci by biotinylated dCas9. Curr Protoc Mol Biol 2018; 123:e64.
28. Thompson A, Schafer J, Kuhn K, *et al.* Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem 2003; 75:1895–1904.
29. Wilson MC, Trakarnsanga K, Heesom KJ, *et al.* Comparison of the proteome of adult and cord erythroid cells, and changes in the proteome following reticulocyte maturation. Mol Cell Proteom 2016; 15:1938–1946.
30. Ashley RJ, Yan H, Wang N, *et al.* Steroid resistance in Diamond Blackfan anemia associates with p57Kip2 dysregulation in erythroid progenitors. J Clin Investig 2020; 130:2097–2110.
31. Narumi R, Shimizu Y, Ukai-Tadenuma M, *et al.* Mass spectrometry-based absolute quantification reveals rhythmic variation of mouse circadian clock proteins. Proc Natl Acad Sci USA 2016; 113:E3461–3467.
32. Simicevic J, Schmid AW, Gilardoni PA, *et al.* Absolute quantification of transcription factors during cellular differentiation using multiplexed targeted proteomics. Nat Methods 2013; 10:570–576.
33. Gillespie MA, Palii CG, Sanchez-Taltavull D, *et al.* Absolute quantification of ■ transcription factors in human erythropoiesis using selected reaction monitoring mass spectrometry. STAR Protoc 2020; 1:100216.
This study provides a detailed protocol on how to perform absolute quantification of key transcription factors and cofactors by targeted mass spectrometry and spike-in internal standards.
34. Takemori N, Takemori A, Tanaka Y, *et al.* MEERCAT: multiplexed efficient cell free expression of recombinant QconCATs for large scale absolute proteome quantification. Mol Cell Proteom 2017; 16:2169–2183.
35. Matsumoto M, Matsuzaki F, Oshikawa K, *et al.* A large-scale targeted proteomics assay resource based on an in vitro human proteome. Nat Methods 2017; 14:251–258.
36. Brun V, Dupuis A, Adrait A, *et al.* Isotope-labeled protein standards: toward absolute quantitative proteomics. Mol Cell Proteom 2007; 6:2139–2149.
37. Singh S, Springer M, Steen J, *et al.* FLEXIQuant: a novel tool for the absolute quantification of proteins, and the simultaneous identification and quantification of potentially modified peptides. J Proteome Res 2009; 8:2201–2210.
38. Zeiler M, Straube WL, Lundberg E, *et al.* A Protein Epitope Signature Tag (PrEST) library allows SILAC-based absolute quantification and multiplexed determination of protein copy numbers in cell lines. Mol Cell Proteom 2012; 11:O111.009613.
39. Yan W, Luo J, Robinson M, *et al.* Index-ion triggered MS2 ion quantification: a novel proteomics approach for reproducible detection and quantification of targeted proteins in complex mixtures. Mol Cell Proteom 2011; 10:M110.005611.
40. Erickson BK, Rose CM, Braun CR, *et al.* A Strategy to combine sample multiplexing with targeted proteomics assays for high-throughput protein signature characterization. Mol Cell 2017; 65:361–370.
41. Ankney JA, Muneer A, Chen X. Relative and absolute quantitation in mass spectrometry-based proteomics. Annu Rev Anal Chem 2018; 11:49–77.
42. Wisniewski JR, Hein MY, Cox J, *et al.* A 'proteomic ruler' for protein copy number and concentration estimation without spike-in standards. Mol Cell Proteom 2014; 13:3497–3506.
43. Hattangadi SM, Martinez-Morilla S, Patterson HC, *et al.* Histones to the cytosol: exportin 7 is essential for normal terminal erythroid nuclear maturation. Blood 2014; 124:1931–1940.
44. Zhao B, Liu H, Mei Y, *et al.* Disruption of erythroid nuclear opening and histone release in myelodysplastic syndromes. Cancer Med 2019; 8:1169–1174.
45. Karayel O, Xu P, Bludau I, *et al.* Integrative proteomics reveals principles of ■ dynamic phosphosignaling networks in human erythropoiesis. Mol Syst Biol 2020; 16:e9813.
This study provides a useful resource for phosphoproteins during erythroid differentiation.
46. Palii CG, Cheng Q, Gillespie MA, *et al.* Single-cell proteomics reveal that quantitative changes in co-expressed lineage-specific transcription factors determine cell fate. Cell Stem Cell 2019; 24:812–820e815.
47. Signer RA, Magee JA, Salic A, *et al.* Haematopoietic stem cells require a highly regulated protein synthesis rate. Nature 2014; 509:49–54.
48. Zaro BW, Noh JJ, Mascetti VL, *et al.* Proteomic analysis of young and old mouse hematopoietic stem cells and their progenitors reveals posttranscriptional regulation in stem cells. eLife 2020; 9:e62210.
49. Yu X, Azzo A, Bilinovich SM, *et al.* Disruption of the MBD2-NuRD complex but not MBD3-NuRD induces high level HbF expression in human erythroid cells. Haematologica 2019; 104:2361–2371.
50. Burgold T, Barber M, Kloet S, *et al.* The Nucleosome Remodelling and Deacetylation complex suppresses transcriptional noise during lineage commitment. EMBO J 2019; 38:e100788.
51. Pessoa Rodrigues C, Herman JS, Herquel B, *et al.* Temporal expression of ■■ MOF acetyltransferase primes transcription factor networks for erythroid fate. Sci Adv 2020; 6:eaaz4815.
This study identified an erythroid-specific role for the histone H4K16 histone acetyltransferase MOF. Furthermore, the authors identified a regulatory network including MOF, RUNX1 and GFI1B that is necessary for erythroid commitment.
52. Leonards K, Almosailleakh M, Tauchmann S, *et al.* Nuclear interacting SET domain protein 1 inactivation impairs GATA1-regulated erythroid differentiation and causes erythroleukemia. Nat Commun 2020; 11:2807.
53. Bludau I, Aebersold R. Proteomic and interactomic insights into the molecular ■ basis of cell functional diversity. Nat Rev Mol Cell Biol 2020; 21:327–340.
This review provides a comprehensive overview of functional diversification in gene expression with a particular focus on the proteome and interactome.
54. Tanimura N, Liao R, Wilson GM, *et al.* GATA/heme multiomics reveals a trace metal-dependent cellular differentiation mechanism. Dev Cell 2018; 46:581–594e584.
55. Vitrinel B, Koh HW, Kar FM, *et al.* Exploiting inter-data relationships in next-generation proteomics analysis. Mol Cell Proteom 2019; 18(8 suppl 1):S5–S14.
56. Cong Y, Motamedchaboki K, Misal SA, *et al.* Ultrasensitive single-cell proteomics workflowidentifies >1000 protein groups per mammaliancell. Chem Sci 2021; 12:1001.
57. Tsai CF, Zhao R, Williams SM, *et al.* An improved boosting to amplify signal ■ with isobaric labeling (iBASIL) strategy for precise quantitative single-cell proteomics. Mol Cell Proteom 2020; 19:828–838.
This study uses nanodroplet-based single cell preparation and isobaric labeling to identify and quantify 1,500 proteins by mass spectrometry in over 100 single cells from leukemia cell lines
58. Specht H, Emmott E, Petelski AA, *et al.* Single-cell proteomic and transcrip- ■■ tomic analysis of macrophage heterogeneity using SCoPE2. Genome Biol 2021; 22:50.
In this study, the authors have used quantitative mass spectrometry to quantify over 3,000 proteins in 1,500 single monocytes and macrophages. They also used the proteome data to build trajectories and discovered a continuous gradient of proteome states for macrophages.
59. Dou M, Zhu Y, Liyu A, *et al.* Nanowell-mediated two-dimensional liquid chromatography enables deep proteome profiling of <1000 mammalian cells. Chem Sci 2018; 9:6944–6951.
60. Woo J, Clair GC, Williams SM, *et al.* Robust, sensitive, and quantitative single-cell proteomics based on ion mobility filtering. bioRxiv (preprint) 2021.