Published in final edited form as:

Transfus Clin Biol. 2010 September ; 17(3): 120–125. doi:10.1016/j.traccli.2010.06.010.

A transcriptome-based examination of blood group expression:

Analyse transcriptomique de l'expression des antigènes de groupe sanguin

S.-J. Noh, Y.T. Lee, C. Byrnes, and J.L. Miller*

Molecular Medicine Branch, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health (NIH), 31 Center Drive, Bethesda, Building 10, Room 9N311, MD 20892, USA

Abstract

Over the last two decades, red cell biologists witnessed a vast expansion of genetic-based information pertaining to blood group antigens and their carrier molecules. Genetic progress has led to a better comprehension of the associated antigens. To assist with studies concerning the integrated regulation and function of blood groups, transcript levels for each of the 36 associated genes were studied. Profiles using mRNA from directly sampled reticulocytes and cultured primary erythroblasts are summarized in this report. Transcriptome profiles suggest a highly regulated pattern of blood group gene expression during erythroid differentiation and ontogeny. Approximately one-third of the blood group carrier genes are transcribed in an erythroid-specific fashion. Low-level and indistinct expression was noted for most of the carbohydrate-associated genes. Methods are now being developed to further explore and manipulate expression of the blood group genes at all stages of human erythropoiesis.

Keywords

Transcriptome; Blood group; Erythropoiesis; RNA

1. Introduction

In the clinic, human red cell antigens provide the structural and immunological basis for hemolytic transfusion reactions [1]. Over the last century, blood agglutination assays were developed to match donor and recipient blood and thereby reduce or prevent related toxicities. The clinical importance of blood antigenicity and agglutination fueled fundamental advances in immunology, hematology and genetic medicine [2]. Initial efforts to specify the nature of erythrocyte antigenicity were conducted through serological testing. Blood from individuals with defined agglutination characteristics was utilized to help identify protein- and carbohydrate-based antigens on the erythrocyte plasma membranes. Meticulous research led to the description of the antigen structures as well as molecules that “carry” each antigen on the erythrocyte surface. It was discovered that many surface molecules carry single antigens, while others encode multiple antigens such as the erythrocyte protein named Band 3, which carries 20 antigens [3]. The ability to predict and

© 2010 Elsevier Masson SAS. All rights reserved.

Corresponding author. jm7f@nih.gov (J.L. Miller).

Conflict of interest statement The authors have no conflict of interest to declare.

Appendix A. Supplementary data Supplementary data associated with this article can be found, in the online version, at doi: 10.1016/j.traccli.2010.06.010.

avoid potentially fatal immune responses through applied research pertaining to erythrocyte antigenicity represents one of the greatest successes of clinical medicine to date.

With the arrival of modern molecular biology, it became possible to expand research in this area to include the blood group genes, and to describe the polymorphisms and mutations responsible for the antigens. The immense clinical infrastructure of transfusion medicine provided investigators with the appropriate research samples, and DNA sequencing technologies provided the ability to locate and carefully map the underlying alleles within the human genome. Over 1000 alleles currently describe the genetic basis for human erythrocyte antigenicity, and genotyping is being pursued as a clinical tool [4]. The digital nature and ease of production of genetic reagents may eventually be ideal for high-throughput applications to blood group phenotyping [5,6]. Based upon the genetic and immunologic information that has already been gathered, blood groups additionally provide fertile soil for research in cellular development and genomic regulation. In this paper, blood group gene expression patterns are described in erythroid and non-erythroid cells.

2. The blood group transcriptome

Classic investigations of erythroid gene transcription focused upon single genes in the context of a particular feature of the erythrocyte. For example, the globin genes were initially studied due to the abundance of those transcripts in reticulocytes [7]. Unfortunately, the abundant globin gene transcripts hampered early investigations of other mRNA species. In the case of blood group antigens, transcriptome analyses were further impeded by the lack of mRNA in the mature erythrocytes of blood donors. During the last two decades, a convergence of cellular and molecular biology provided the means to overcome these barriers and generate full transcriptome profiles from erythroblasts prior to terminal differentiation. Cell culture techniques permitted the identification and isolation of erythroblasts at all stages of differentiation [8]. Furthermore, flow cytometry enhanced the isolation of high-purity erythroblast populations [9]. Gene sequencing, PCR and other improvements in molecular biology led to the robust quantitation of mRNA species. Finally, with the advent of high-throughput technologies, the generation of entire transcriptome profiles became possible. New bioinformatic tools have allowed investigators to conduct research analyses via the Internet. With bioinformatics, it became apparent that expressed sequence tags (ESTs) could be used in lieu of full-length transcripts to describe the erythroid transcriptome [10]. Those ESTs were mapped into the human genome for identification of the encoded gene and comparison with related sequences. The gene named ART4 that encodes the Dombrock carrier molecule was originally identified using this approach in the late 1990s [11].

After the human genome was mapped, high throughput sequencing was no longer required for generation of transcriptome profiles. Instead, mRNA species and their derivatives were hybridized to arrays of complementary nucleic acids for detection and quantitation. While the ability to discover truly novel mRNA species is lost with many arrays, the RNA “chips” are adequate for profiling the expression patterns of known or predicted sequences [12]. Since the genes encoding the human blood group carrier molecules and the related sugar-modifying enzymes have become known, array-based technologies are useful for exploring the expression of those genes. Specialized arrays may be produced for mRNA profiling [13]. Alternatively, bioinformatic software is available to examine subsets of transcripts within the erythroid transcriptome. Similar technologies have broadened the scope to encompass the examination of polymorphisms within the human genome that encode each allele for the blood group antigens. It is predicted that informatic analyses of whole-genome and whole-transcriptome sequencing will eventually replace the existing technologies to provide more robust diagnostic tools. In the clinic, one can envision a vast genetic web of information for

each patient that may be accessed for personalized clinical care. Ultimately, genetic information pertaining to the blood groups should be useful for clinical hematology, transfusion medicine and infectious disease. Due to non-erythroid expression of some blood groups, the same information may eventually be helpful in other medical areas.

3. Patterning blood group gene expression

In our laboratory, array technologies replaced high throughput sequencing as a means of describing the erythroid transcriptome. The Affymetrix platform replaced custom arrays due to their reproducibility, ease of analysis and the ability to compare separate experiments [14]. Importantly, it is emphasized here that the cellular phenotype generated from mRNA is not equivalent to a protein-, lipid- or carbohydrate-based phenotype.

Representative blood group transcriptomes from differentiating human erythroblasts are shown in Fig. 1. To generate the erythroblast transcript profiles, the cells were isolated in high purity [14,15]. Erythroblasts were isolated in purity using flow cytometry from ex vivo cultures of CD34 cells prior to, and after expression of glycophorin A on their plasma membranes. As shown, the expression of two-thirds of the genes was increased with maturation of the cells as determined by glycophorin A on the cell surface (Table 1). Interestingly, a relatively high-level signal for GYPA mRNA was detected before the protein was present on the cell surface. Also, mRNA encoding the ABO glycosyltransferases was negligible in these committed erythroblast populations. While consistent with the results of others, the low-level expression of ABO in committed erythroid cells is not well understood [16].

Transcriptome profiles were additionally generated directly from reticulocyte mRNA contained in circulating umbilical cord and adult blood (Fig. 1, Table 1). Blood from term pregnancies and healthy adult donors were used to generate the normalized intensities. When these studies began, it was predicted that reticulocyte mRNA levels would roughly match the antigen levels on the mature cells. According to published reports, antigens associated with the LU, YT, ART4, XG, CR1 and BSG genes are increased in adult blood, and the LW antigens are reportedly decreased [3]. In comparison with the antigen levels, the levels of mRNA encoding the carrier molecules were equivalent or slightly increased in the adult reticulocyte populations. On the other hand, very distinct decreases in SLC4A1, XK, SLC14A1, GYPA, GYPB, GYPE, BSG, RHCE and RHD mRNA species were detected in the adult reticulocytes (Table 1). The differences in cord and adult blood levels for these genes were consistent between the donor subject pairs (data not shown). Since there is no evidence for similarly decreased levels of these antigens on the mature cell surface during ontogeny, other causes for the decreased mRNA levels compared to their protein relatives are being considered. The stability of those transcripts in reticulocytes or post-transcriptional mechanisms for the regulation of membrane proteins during terminal erythroid maturation may be involved.

4. Comparing blood group transcriptomes among tissues

Transcriptome profiles from a variety of cells and tissues are now available in the public domain, including those describing erythroblasts and reticulocytes. Fig. 2 shows a comparison of a sampling of blood group genes that are provided by BioGPS [17]. A complete description of the blood group gene profiles is shown in Supplementary data, Fig. 1. Approximately 50% of the genes demonstrated the highest level of mRNA detection among erythroid tissues. Within that group, erythroid-focused expression was seen for 13 genes (ART4, ERMAP, FUT1, GYPA, GYPB, GYPE, ICAM4, KEL, RHCE, RHD, SLC4A1, XG and XK). The similarity of expression among this group of 13 erythroid-focused genes suggests distinct mechanisms for transcriptional regulation in erythroid versus

nonerythroid tissues. However, erythroblast maturation related differences in mRNA and protein expression patterns support the notion that several levels of erythroid-specific gene regulation exist during erythropoiesis (Fig. 1) [18].

A separate group of 15 genes are expressed in multiple tissues at variable levels. Six of those genes are expressed predominantly in erythroid cells among the representative tissues (ACHE, BSG, C1GALT1, DARC, GYPC and SLC14A1). Seven genes (CD44, CD55, CD99, CD151, AQP1, AQP3 and C1GALT1C1) showed a broad range of expression with erythroid expression not being predominant. Expression of two genes was specific for non-erythroid tissues. The C4 gene transcripts were detected primarily in the liver, and BCAM in the thyroid. Remarkably, the profiles of the remaining seven genes (ABO, A4GALT, B3GALT3, CR1, GCNT2, FUT2, FUT3, SEMA7A) showed low-level expression in most human tissues. The group of genes with low-levels of transcripts contained the majority of the carbohydrate-related blood groups. The grouped expression of the genes may be associated with selective pressures that presumably contributed to antigen formation.

5. Needed: better experimental models

The functional importance of most erythrocyte antigens is unknown. Based upon the discovery of DARC as a malarial receptor, it is logical to conclude that several antigens may have arisen in response to malaria [19]. For other antigens, separate infections or diseases may be involved [20]. To further explore this interesting topic, improved experimental models need to be developed. In the absence of an appropriate erythroid stress, our ability to determine any biological role for the antigens is suboptimal. Cell culturing under defined conditions may be helpful for determining the regulation of blood group genes and the function of the gene products during normal and stressed erythropoiesis. Major advances in cell culture methods for studying erythropoiesis using primary erythroid cells have already been made. Successful enucleation of the erythroblasts in culture with production of mature erythrocytes seems to be an important goal. Cytokine cocktails, fibronectin and co-culture with feeder cells have been used to achieve this goal [21-23]. Flow cytometry is becoming an essential tool for identifying and isolating the cells and their components during the final stages of maturation (Fig. 3). For manipulation of the blood group genes, lentiviral transduction can be incorporated into the primary cell culture systems (Fig. 4) [24]. Ultimately, the ability to more fully explore and modify erythropoiesis with engineered or donor-defined cells should greatly assist in these efforts.

6. Summary: the bright future ahead

Over the last two decades, studies of blood group genes and their products have transitioned from a discovery phase toward high-throughput patterning of the entire group. In the context of the transcriptome, it is now possible to obtain profiles directly from clinical samples. Examination of those profiles suggests that erythroid cells possess distinct mRNA phenotypes for the blood group genes that are not fully predicted by the associated antigen expression levels on the surface of the mature cells. Transcriptome profile comparisons suggest that only half of the blood group carrier genes are transcribed predominantly in erythroid cells. Erythroid specificity, or the lack thereof, may be useful for understanding the formation and function of antigen alleles. Studies are now underway to develop sufficiently robust methods for exploring the regulation of blood group gene expression and function.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The Intramural Research Program of the NIH, NIDDK supported this research.

References

1. Davenport RD. Pathophysiology of hemolytic transfusion reactions. *Semin Hematol.* 2005; 42:165–8. [PubMed: 16041666]
2. Landsteiner K. The Nobel lectures in immunology. The Nobel Prize for Physiology or Medicine, 1930, awarded to Karl Landsteiner In recognition of his discovery of human blood groups. *Scand J Immunol.* 1990; 32:1–12. [PubMed: 2200106]
3. Daniels GL, Anstee DJ, Cartron JP, Dahr W, Fletcher A, Garratty G, et al. International society of blood transfusion working party on terminology for red cell surface antigens. *Vox Sang.* 2001; 80:193–7. [PubMed: 11449960]
4. Avent ND, Martinez A, Flegel WA, Olsson ML, Scott ML, Nogués N, et al. The BloodGen project: toward mass-scale comprehensive genotyping of blood donors in the European Union and beyond. *Transfusion.* 2007; 47(Suppl. 1):40S–146S. [PubMed: 17593285]
5. Araujo F. Real-time PCR assays for high-throughput blood group genotyping. *Methods Mol Biol.* 2009; 496:25–37. [PubMed: 18839102]
6. Veldhuisen B, van der Schoot CE, de Haas M. Blood group genotyping: from patient to high-throughput donor screening. *Vox Sang.* 2009; 97:198–206. [PubMed: 19548962]
7. Williamson R. The isolation and properties of globin messenger RNA. *Biochimie.* 1972; 54:597–9. [PubMed: 4654156]
8. Wojda U, Noel P, Miller JL. Fetal and adult hemoglobin production during adult erythropoiesis: coordinate expression correlates with cell proliferation. *Blood.* 2002; 99:3005–13. [PubMed: 11929793]
9. Miller JL, Njoroge JM, Gubin AN, Rodgers GP. Prospective identification of erythroid elements in cultured peripheral blood. *Exp Hematol.* 1999; 27:624–9. [PubMed: 10210320]
10. Gubin AN, Njoroge JM, Bouffard GG, Miller JL. Gene expression in proliferating human erythroid cells. *Genomics.* 1999; 59:168–77. [PubMed: 10409428]
11. Gubin AN, Njoroge JM, Wojda U, Pack SD, Rios M, Reid ME, et al. Identification of the dombrock blood group glycoprotein as a polymorphic member of the ADP-ribosyltransferase gene family. *Blood.* 2000; 96:2621–7. [PubMed: 11001920]
12. McGall GH, Christians FC. High-density genechip oligonucleotide probe arrays. *Adv Biochem Eng Biotechnol.* 2002; 77:21–42. [PubMed: 12227735]
13. Lee YT, Miller LD, Gubin AN, Makhlof F, Wojda U, Barrett AJ, et al. Transcription patterning of uncoupled proliferation and differentiation in myelodysplastic bone marrow with erythroid-focused arrays. *Blood.* 2001; 98:1914–21. [PubMed: 11535529]
14. Goh SH, Josleyn M, Lee YT, Danner RL, Gherman RB, Cam MC, et al. The human reticulocyte transcriptome. *Physiol Genomics.* 2007; 30:172–8. [PubMed: 17405831]
15. Tanno T, Bhanu NV, Oneal PA, Goh SH, Staker P, Lee YT, et al. High levels of GDF15 in thalassemia suppress expression of the iron regulatory protein hepcidin. *Nat Med.* 2007; 13:1096–101. [PubMed: 17721544]
16. Edvardsson L, Dykes J, Olsson ML, Olofsson T. Clonogenicity, gene expression and phenotype during neutrophil versus erythroid differentiation of cytokine-stimulated CD34+ human marrow cells in vitro. *Br J Haematol.* 2004; 127:451–63. [PubMed: 15521924]
17. Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, et al. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.* 2009; 10:R130. [PubMed: 19919682]
18. Southcott MJ, Tanner MJ, Anstee DJ. The expression of human blood group antigens during erythropoiesis in a cell culture system. *Blood.* 1999; 93:4425–35. [PubMed: 10361141]
19. Miller LH, Mason SJ, Dvorak JA, McGinniss MH, Rothman IK. Erythrocyte receptors for (*Plasmodium knowlesi*) malaria: Duffy blood group determinants. *Science.* 1975; 189:561–3. [PubMed: 1145213]

20. Anstee DJ. The relationship between blood groups and disease. *Blood*. 2010; 115(23):4635–43. [PubMed: 20308598]
21. Muta K, Krantz SB, Bondurant MC, Wickrema A. Distinct roles of erythropoietin, insulin-like growth factor I, and stem cell factor in the development of erythroid progenitor cells. *J Clin Invest*. 1994; 94:34–43. [PubMed: 7518834]
22. Patel VP, Lodish HF. A fibronectin matrix is required for differentiation of murine erythroleukemia cells into reticulocytes. *J Cell Biol*. 1987; 105:3105–18. [PubMed: 2961771]
23. Douay L. Transfusional perspectives for ex vivo control of hematopoiesis. *Transfus Clin Biol*. 2003; 10:151–5. [PubMed: 12798848]
24. Hanawa H, Persons DA, Nienhuis AW. High-level erythroid lineage-directed gene expression using globin gene regulatory elements after lentiviral vector-mediated gene transfer into primitive human and murine hematopoietic cells. *Hum Gene Ther*. 2002; 13:2007–16. [PubMed: 12489996]

Fig. 1.

Expression patterns of the blood group genes in human erythroid tissues. Transcriptome profiles were generated from human erythroid cells as previously described [14,15]. Among the transcriptome data, probes were identified that hybridize blood group genes as denoted in Table 1. Left panel: probe intensity comparison using mRNA from CD34 cultured erythroblasts. Transferrin receptor (CD71) and glycophorin A (GPA) phenotyping were used for cell identification and flow cytometry sorting. The erythroblast populations were sorted according to their maturation stage [CD71(+) GPA (-) denotes the less mature phenotype]. Right panel: probe intensity comparison using mRNA from purified reticulocytes in umbilical cord ($n = 6$) and adult ($n = 6$) blood. All clinical samples were obtained with donor consent and NIH clinical research approval that comply with the ethical recommendations in the Helsinki declaration.

Fig. 2. Representative BioGPS profiles of erythroid and non-erythroid expression of the blood group genes. Normalized gene expression profiles were organized according to cell and tissue type as shown on the y-axis. Expression in erythroblasts is shown as a red bar for comparison. ART4 and KEL patterns are provided as examples of erythroid-focused expression. DARC and CD44 represent multi-tissue expression profiles with major and minor erythroid intensities. ABO represents a pattern of low-level, pan-tissue expression seen for most of the carbohydrate-associated genes. Profiles for the remaining blood group genes are shown in Supplementary data Fig. 1.

Fig. 3. Flow cytometry of cultured erythroid cells undergoing terminal maturation. Analysis of thiazole orange stained samples is shown to demonstrate the sorting strategy for extruded nuclei, orthochromatic normoblasts and mature cells. The sorted populations are shown on the right with arrows pointing to populations identified by flow cytometry.

Fig. 4. Mature cultured erythroid cells transduced with GFP-producing lentivirus. A nucleated and enucleated cell are shown.

Table 1

Affymetrix HG-U133 Plus 2.0 GeneChip probe normalized intensities of the blood group genes.

Probe	Gene name	Blood group	Sorted erythroblasts		Reticulocytes	
			GPA (-)	GPA (+)	Cord blood	Adult blood
214504_at	ABO	ABO	0.2	0.2	0.6	0.9
214428_x_at	C4A/C4B	Chido Rodgers	7.3	2.9	7.4	10.0
207542_s_at	AQP1	Colton	10.45	75.1	3.2	4.8
201926_s_at	CD55	Cromer	24.4	20.4	4.0	4.9
1552713_a_at	SLC4A1	Diego	9.2	120.5	39.7	21.9
207220_at	ART4	Dombrock	2.3	12.6	1.3	1.3
208335_s_at	DARC	Duffy	9.2	32.8	1.8	2.8
202947_s_at	GYPE	Gerbich	176.6	304.5	229.6	303.3
203747_at	AQP3	Gill	5.1	12.3	3.6	2.3
206109_at	FUT1	H	3.1	10.8	2.3	2.5
210608_s_at	FUT2	H	2.0	2.1	2.5	4.2
211020_at	GCNT2	I	3.3	7.3	0.2	0.5
204489_s_at	CD44	Indian	18.4	16.3	12.2	14.5
210083_at	SEMA7A	JMH	0.3	0.3	1.8	2.4
206077_at	KEL	Kell	40.7	84.5	5.4	6.4
206698_at	XK	Kell	90.0	149.0	69.7	11.0
205856_at	SLC14A1	Kidd	1.6	16.0	10.3	3.9
217484_at	CR1	Knops	1.7	2.0	2.3	2.3
207194_s_at	ICAM4/CD242	LW	20.6	31.3	2.5	2.8
216010_x_at	FUT3	Lewis	0.2	0.9	3.1	2.9
40093_at	BCAM	Lutheran	0.3	0.7	3.6	4.8
211821_x_at	GYPE	MNS	116.8	403.6	107.4	14.7
207459_x_at	GYPB	MNS	78.0	268.5	115.0	18.9
207854_at	GYPE	MNS	15.0	56.1	13.1	2.1
208677_s_at	BSG/CD147	OK	72.7	165.2	122.1	82.1
219488_at	A4GALT	P-related	9.0	6.9	1.3	1.3
211812_s_at	B3GALT3	P-related	29.1	16.0	0.3	0.6
204306_s_at	CD151	RAPH	4.5	2.1	0.9	0.7

Probe	Gene name	Blood group	Sorted erythroblasts		Reticulocytes	
			GPA (−)	GPA (+)	Cord blood	Adult blood
216317_x_at	RHCE	RH	54.9	300.1	20.0	5.6
210430_x_at	RHD	RH	27.1	160.0	6.2	1.2
219905_at	ERMAP	Scianna	25.1	69.3	1.8	1.1
219439_at	CIGALT1	T/Th	8.5	20.1	2.1	1.5
219283_at	CIGALT1C1	T/Th	21.1	15.1	0.5	1.2
201029_s_at	MIC2/CD99	XG	104.7	68.0	5.2	5.2
1554062_at	XG	XG	0.1	1.6	2.2	2.0
205377_s_at	ACHE	YT	2.1	9.1	0.6	0.9