Biological and Genetic Determinants of Red Blood Cell Glycolysis

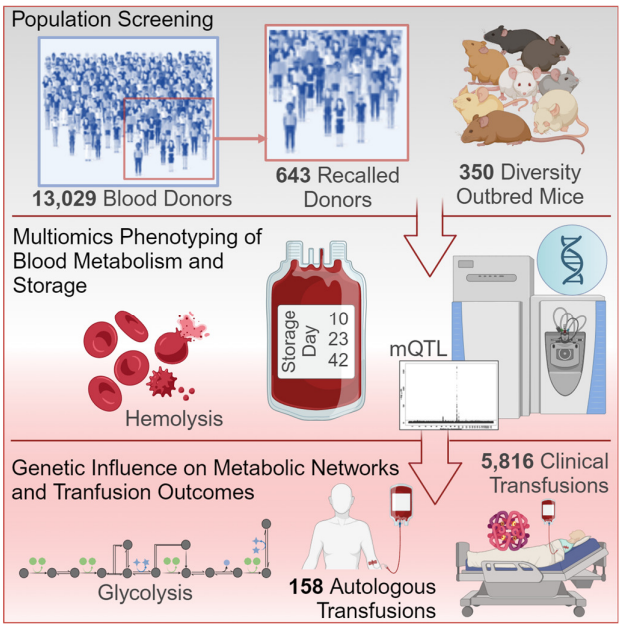
Travis Nemkov,[1,2] Daniel Stephenson,[1] Eric Earley,[3] Gregory R Keele,[4] Ariel Hay,[5] Christopher Erickson,[1] Monika Dzieciatkowska,[1] Alicia Key,[1] Amy Moore,[3] Mars Stone,[6,7] Xutao Deng, [6,7] Steven Kleinman,[8] Steven L Spitalnik,[9] Eldad A Hod,[9] Kirk C Hansen,[1,2] Gary A Churchill,[4] Nareg Roubinian,[6,7,10] Philip J. Norris,[6,7] Michael P. Busch,[6,7] James C Zimring,[5] Grier P. Page,[3] Angelo D'Alessandro[1,2*]

1) Department of Biochemistry and Molecular Genetics, University of Colorado Denver – Anschutz Medical Campus, Aurora, CO, USA;
2) Omix Technologies Inc, Aurora, CO, USA;
3) RTI International, Atlanta, GA, USA;
4) Jackson lab, Bar Harbor, ME, USA;
5) Department of Pathology, University of Virginia, Charlottesville, VA, USA;
6) Vitalant Research Institute, San Francisco CA, USA;
7) Department of Laboratory Medicine, University of California San Francisco, CA, USA
8) University of British Columbia, Victoria, British Columbia, Canada;
9) Department of Pathology, Columbia University Irving Medical Center, New York, NY, USA;
10) Kaiser Permanente Northern California Division of Research, Oakland, CA

***Corresponding author:**

Angelo D'Alessandro, PhD
Department of Biochemistry and Molecular Genetics
University of Colorado Anschutz Medical Campus
12801 East 17th Ave., Aurora, CO 80045
Phone # 303-724-0096
E-mail: angelo.dalessandro@cuanschutz.edu

Graphical abstract



Abstract

Glycolysis is a central metabolic pathway in health and disease,[1] an essential one in mature red blood cells (RBCs), which rely on the Embden-Meyerhof-Parnas pathway, i.e., glycolysis, as the sole source of energy generation.[2] During aging in vivo and in vitro under blood bank conditions,[3] the adenosine triphosphate (ATP) generated via glycolysis in mitochondria-devoid mature RBCs is essential for the modulation of oxygen kinetics,[4] deformability[2] and, ultimately, intra- and extra-vascular hemolysis.[3] Here we draw upon a cohort of 13,029 volunteers from the Recipient Epidemiology and Donor Evaluation Study (REDS) to identify biological and genetic traits associated with heterogeneity in glycolytic phenotypes. Age, sex and ethnicity influenced glycolysis, with additional effects due to genetic polymorphisms in regions coding for rate-limiting enzymes of glycolysis,[1] phosphofructokinase 1 (PFKP) and hexokinase 1 (HK1), and for the ADP-ribosyl cyclase CD38. The gene-metabolite associations were validated by testing glycolysis in fresh and stored RBCs from 350 Diversity Outbred mice.[5] These findings were further corroborated via multi-omics analyses of 1,929 samples at storage day 10, 23 and 42 in 643 RBC units from donors who were selected amongst the index cohort based on their extreme hemolytic propensity. Finally, we show that glycolysis, ATP levels, breakdown and deamination into hypoxanthine are associated with increased hemolytic propensity and vesiculation *ex vivo*, as well as with extravascular hemolysis in vivo, both in healthy autologous transfusion recipients from the Donor Iron Deficiency Study and in 4,700 heterologous critically ill patients receiving blood products from the REDS donors. This work advances our understanding of genetic and non-genetic factors that regulate glycolysis in mature RBCs, a simplified model of mammalian cell metabolism,[6] a cell type where dysregulated glycolysis holds broader implications in hemolytic disorders, offering potential avenues for novel therapeutic interventions and transfusion strategies.

Main

A culmination of a near century of study involving many pioneers in the burgeoning field of biochemistry,[7] the Embden-Meyerhof-Parnas pathway or glycolysis explained the cell's ability to convert glucose into adenosine triphosphate (ATP) even in the absence of oxygen.[8] Altough the details of a complex interplay between glycolysis and mitochondrial metabolism have since been delineated, certain cell types such as red blood cells (RBCs) are devoid of organellar components and, therefore, are completely reliant on glycolysis both for energy generation and as a central network hub for metabolic homeostasis. RBCs comprise ~83% of all cells in the human body.[9] Ease of collection and shear abundance have made RBCs an attractive model for investigating metabolism in the early days of biochemistry.[10] Unlike other vertebrates,[11] mammalian RBCs evolved to maximize oxygen carrying capacity by increasing hemoglobin content per cell, a feature achieved by discarding nuclei and organelles. Given their lack of de novo protein synthesis capacity, RBCs rely on dynamic metabolic reprogramming in response to environmental stimuli, including hypoxia or oxidant stress (e.g., at altitude or during exercise).[3] For example, small molecule metabolites such as ATP play critical roles in hemoglobin allostery by stabilizing the tense deoxygenated state and promoting oxygen release to tissues.[2] Relevant to biology at large, mitochondria-devoid RBCs represent a simplified model of eukaryotic cell metabolism, a cell type that exclusively relies on glycolysis as the main source of energy for the synthesis of ATP. For the reasons above, it is not a coincidence that the earliest efforts to reconstruct human metabolic networks in silico have been centered on RBC metabolism.[6]

In addition to the central role RBCs play in mediating systemic gas and metabolite exchange throughout the body, understanding erythrocyte metabolism holds the key to optimizing storage conditions and predicting post-transfusion performance of blood products. Transfusion of packed RBCs is the most common in-hospital medical procedure after vaccination, a life-saving intervention for tens of millions of civilian and military patients whose survival depends solely on chronic or massive transfusions, with no current alternative medical interventions. Altered energy metabolism is a hallmark of the so-called storage lesion, a series of biochemical and morphological alterations that ultimately increase the propensity of stored RBCs to hemolyze ex vivo in the refrigerated unit or in vivo upon transfusion.[12] Storage-dependent depletion of ATP pools, consumption of glucose and accumulation of lactate are key biomarkers of the metabolic storage lesion.[12] Although the clinical relevance of the storage lesion remains unresolved, it is now clear that RBC metabolism under refrigerated storage conditions is impacted by storage temperatures, additive solutions, donor dietary or other exposures (e.g., smoking, alcohol consumption, drugs that are not grounds for blood donor deferral), and also by factors such as donor age, sex, ethnicity and body mass index (BMI).[13] In other

words, how RBCs store and how they perform upon transfusion are a function of donor genetics[14] and non-genetic factors.[15]

The recent implementation of high-throughput and cost-effective genomics and metabolomics platforms has enabled powerful large scale metabolic Quantitative Trait Loci (mQTL) studies.[16] This new body of work has expanded Garrod's classic concept of chemical individuality,[17] through comprehensive genome-wide association studies that reveal how genetic variation contributes to heterogeneity in blood metabolite levels.[18] Most of these studies have interrogated the genetic underpinnings of pathological metabolic conditions. These efforts underlie the foundation of modern day clinical chemistry, especially the focus on inborn errors of metabolism, such as those linked to hemolytic disorders (e.g., pyruvate kinase deficiency or glucose 6-phosphate dehydrogenase deficiency).[3] Despite the importance of metabolism and glycolysis in health and disease, and decades of research on the biochemical determinants of glycolytic fluxes in mammals,[1] the genetic and non-genetic underpinnings of the biochemical heterogeneity in humans and mice are incompletely understood.

To bridge this gap, we performed high-throughput metabolomics analyses on leukocyte-filtered pRBCs donated by 13,029 volunteers at four different blood centers across the United States and stored for 42 days (i.e., at the end of the FDA-approved shelf-life - **Figure 1.A**). The donor population was balanced with respect to sex (6,507 males and 6,522 females) and blood center (from a minimum of 3,019 to a maximum of 3,564 per center). The age range spanned from 18 to 87 (median age 47). Body mass index (BMI) ranged from 10-78 (median 26.6). The majority of donors in this cohort were Caucasians (7,037), with substantial representation of individuals of Asian (1,602), African (1,542) and Hispanic descent (1,153). Additive solution 1 (AS-1) and AS-3 were the only two additives used.

Donor sex, age, ethnicity and BMI, as well as latent variables impact RBC metabolism of stored RBCs

Linear discriminant analysis (LDA) of all the samples as a function of donor age, sex, body mass index (BMI) and ethnicity identified strong effects of each factor on the metabolic heterogeneity of end of storage pRBC units (**Supplementary Figure 1.A-E**). Network analysis of metabolite-metabolite correlations across the whole population revealed a central core of metabolites involved in glycolysis (**Supplementary Figure 1.H**), significantly associated to donor demographics (**Supplementary Figure 1.I**), including ATP and lactate (**Figure 1.B**). K means clustering was used to identify latent factors that influence metabolic patterns of end of storage units beyond demographics (**Figure 1.C**). Hierarchical clustering analysis of samples based on the top 50 significant metabolites that discriminate the 10 clusters by ANOVA detected a strong signal

separating donors in cluster 4 from the rest of the donors in this study (**Supplementary Figure 1.F**). Independent network analyses based on metabolite-metabolite Spearman correlations were performed for each one of the 10 clusters derived from the K means analysis (**1.D**). Results indicated a metabolic rewiring in glycolysis for samples in cluster 4, which were characterized by the highest levels of ATP (**Figure 1.E**) and concomitantly elevated levels of all glycolytic metabolites (heat map by ATP levels, 5th vs 95th percentile in **Figure 1.F;** correlations to ATP in **Supplementary Figure 1.G-L**). Higher ATP levels in units of donors in cluster 4 were not explained by differences in donor demographics (e.g., age - **Figure 1.G**) or storage additives. Donors with the highest levels of ATP and lowest levels of hypoxanthine – a breakdown and deamination product of ATP – showed markedly distinct end of storage metabolic phenotypes (**Supplementary Figure 2.A-C**).

Glycolysis is influenced by donor sex, age, ethnicity and BMI

In light of the analyses above, first we analyzed end of storage heterogeneity in RBC glycolysis across donors based on demographics to determine the impact on this pathway from factors like age, alone or in combination with sex, ethnicity and BMI (**Supplementary Figure 3**), Results clearly show a progressive decline as a function of donor age in the levels of ATP and all glycolytic intermediates, but increases in lactate (**Supplementary Figure 3.A**). This phenomenon showed sex dimorphism, with lower levels of glucose and higher levels of glycolytic metabolites in males across the age range (e.g., 2,3-bisphosphoglycerate, phosphoenolpyruvate and lactate, but not ATP). Altogether, these results are suggestive of higher glycolytic rates in males to fuel the Rapoport-Luebering shunt without net advantage on the generation of larger RBC ATP pools (**Supplementary Figure 3.B**). Of note, sex dimorphism in glucose metabolism – as inferred from steady state pools of glycolytic metabolites - was not observed for donors past 50 years of age, the average national age for menopause in the United States.[19]

RBCs from donors with the highest BMIs (morbidly obese; >40) were characterized by significantly higher levels of lactate compared to all the other BMI groups within the age window of 30-60 (**Supplementary Figure 3.C**).

Ethnicity measurably affected glycolysis, with end-of-storage pRBCs from donors with Hispanic or other/multiple ethnic ancestry being characterized by higher ATP and 2,3-DPG, and of all glycolytic intermediates through pyruvate, but not lactate; the latter was highest in donors of Caucasian or East Asian descent (**Figure Supplementary Figure 3.D**).

Genetic polymorphisms associated with the glycolytic heterogeneity of end of storage units from 13,029 REDS RBC Omics Index donors

The observed impact of ethnicity on ATP and glycolysis in the REDS RBC Omics Index donor population suggested that certain heritable genetic traits contribute to heterogeneity in end of storage levels of these metabolites. In addition to metabolomics analyses, we tested donors for ~879,000 single nucleotide polymorphisms (SNPs), which enabled mapping of metabolite Quantitative Trait Loci (mQTL, **Figure 2.A**). A summary Manhattan plot of mQTL is shown for all glycolytic metabolites merged (**Figure 2.B**), or just for ATP (**Supplementary Figure 4**). Manhattan plots are shown for lactate (**Figure 2.B**) and each of the glycolytic metabolites (**Supplementary Figure 5**). Related Locus Zoom plots for the most common SNPs in phosphofructokinase P (PFKP, chromosome 10) that associate with ATP and lactate (FDR-corrected p-value = e-163) are shown in **Supplementary Figure 6.** Of note, SNPs in PFKP also ranked as the strongest factors influencing hypoxanthine levels, followed by SNPs in the region coding for adenosine deaminase (ADA, chromosome 12 – **Figure 2.C**).

This analysis also identified multiple common hot spot areas in the regions coding for other rate-limiting enzymes of glycolysis, hexokinase 1 (HK1, chromosome 10), phosphoglucomutase 2 like 1 (PGM2L1, chromosome 11), pyruvate kinase (PKLR, chromosome 1 – **Supplementary Figure 5**). Similarly, end of storage levels of pyruvate showed a strong association with the rate-limiting enzyme of the pentose phosphate pathway, glucose 6-phosphate dehydrogenase (G6PD – **Supplementary Figure 5**). For hexose phosphate, an extremely significant hit (p<e-300) was observed on fructosamine 3-kinase (FN3K), a deglycating enzyme and modulator of NRF2 activity in cancer,[20] whose polymorphisms have been implicated in the pathogenesis of diabetic complications.[21] For glycolytic metabolites downstream to glyceraldehyde 3-phosphate, a common hit absent in the hexose phosphate group was the ectonucleotidase CD38 or the neighboring bone marrow stromal antigen - BST1 (also known as CD157), an ectonucleotidase with similar pH-dependent NAD(P)+ hydrolase or ADP-ribosyl cyclase activity to CD38[22,23], both on chromosome 4 (**Supplementary Figure 5**). For glyceraldehyde 3-phosphate and ribose phosphate, a hit on RHCE (chromosome 1) suggests a potential cross-talk between the RH group components and the metabolic steps linking the non-oxidative branch of the pentose phosphate pathway to triose phosphate compounds (**Supplementary Figure 5**). HSB1L-MYB (chromosome 6) and CPPED1 (chromosome 16) identifies an association between ATP and phosphoglycerate level and hemoglobin variants, including hemoglobin F[24] and alpha-thalassemia in ATR-16 syndrome, deletions of ~177 to ~2000 kb affecting one allele of the well-characterized, gene dense, distal region of chromosome 16 (16p13.3).[25] A summary overview of the gene-metabolite associations is shown in **Figure 2.D**.

Interactions between PFKP polymorphisms and donor age, sex and ethnicity in index and recalled human donor cohorts

Correlation analysis of PFKP rs2388595 alleles across donors confirms strong positive associations with lactate and negative association with hypoxanthine (**Supplementary Figure 7.A**), overall impacting all glycolytic metabolites (**Supplementary Figure 7.B**). These findings were validated in the recalled donor cohort (**Figure 2.E**), whereby PFKP protein levels - as quantified by proteomics – were identified as the top positive correlates to the PFKP rs2388595 SNP, suggestive that this intronic variant does impact protein expression.

Frequency of PFKP rs2388595 homozygosity was highest in donors of Caucasian (25%) or Asian descent (23%), and lowest in donors of Hispanic descent (18% - **Figure 2.F**). Effects of biological variables like donors age and sex were further exacerbated by PFKP SNPs, especially with respect to glucose and fructose 1,6-bisphosphate consumption and lactate accumulation, with concomitantly significantly lower levels of hypoxanthine in older male donors who carry two alleles of the PFKP rs2388595 SNP (**Figure 2.G-H; Supplementary Figure 7.D-E**).

Consistent with a donor-dependency of glycolysis (at the biological and genetics level), the steady state abundance of glycolytic metabolites at storage day 42 was significantly reproducible across multiple donations from the same donors who were enrolled in both the index and recall phase of the REDS RBC Omics study (**Supplementary Figure 7.C**). Consistent with the data from the index phase, donor genetics had a compounded effect on glycolysis and hypoxanthine as a function of storage duration and PFKP SNPs in recalled donor units (**Supplementary Figure 7.F**).

Genetic underpinnings of glycolytic heterogeneity in fresh and stored murine RBCs in the J:DO study and validation in humans from the index and recalled cohorts

To validate and expand on the human mQTL findings, we measured metabolite levels in fresh and stored murine RBCs from 350 mice from the Jackson Laboratory's Diversity Outbred (J:DO) mouse population. Murine RBCs were stored for 7 days, which was chosen upon pilot studies to identify the optimal storage range resulting in end of storage PTR comparable to humans,[26] while also accounting for the wide variability in PTR amongst the J:DO mice. Briefly, eight genetically diverse founder strains were crossed for >40 generations to establish this cohort with highly recombined genomes (**Figure 3.A**). These mice were genotyped for 137,192 loci, and fresh and stored RBCs were tested for the levels of glycolytic metabolites (**Figure 3.A**). Results identified a hot spot region of co-mapping mQTLs on murine chromosome 5 for ATP levels and glycolytic metabolites (**Figure 3.B**), as highlighted by Manhattan plot, hive plot and heat map illustrations of the results (**Figure 3.C-E**). A representative Manhattan plot for ATP is shown in **Figure 3.F**, and shows the association with a

polymorphic region on chromosome 5 (**Figure 3.G**) which encodes CD38 (**Figure 3.H**). This very same region associates with all glycolytic and pentose phosphate pathway (PPP) metabolites in stored and fresh murine RBCs (**Figure 3.D-E;** annotated versions in **Supplementary Figure 8**).

Interestingly, the CD38 findings for metabolites downstream to glyceraldehyde 3-phosphate in glycolysis and pentose phosphate pathway metabolites were consistent between the J:DO mice and humans (**Supplementary Figure 9** shows two representative side-by-side comparisons). As a validation of the mouse work in humans, the most significant CD38 SNP rs4478185 in the REDS RBC Omics Index cohort, while rare (homozygous recessive in 0.2% and heterozygous in 3.4% of the index population - **Supplementary Figure 10.A**), was significantly associated with the levels of glycolytic metabolites and ATP (**Figure 3.I**). As a reassuring internal validation of the quality of the metabolomics data, we observed a positive association with NAD(P)+ and negative association of this SNP with nicotinamide and ADP-ribose, substrates and products of the enzymatic activity of CD38, respectively (**Figure 3.I**). Indeed, mQTL analysis for ADP-ribose highlighted both CD38/BST1 and PFKP SNPs as top loci associated with the end of storage levels of this metabolite in the index donor population (**Supplementary Figure 13**). These observations held true also in the recalled donor cohort, a subset of 1,929 longitudinal samples from 643 donors (**Supplementary Figure 12**). Interestingly, in the multi-omics analysis of the recalled donor population, the CD38 SNP rs4478185 was negatively associated with osmotic hemolysis, as well as with the levels of the atypical chemokine receptor 1 (ACKR1), one of the components of the Duffy blood group (**Supplementary Figure 5.I**).

While shared metabolite-gene associations in CD38 were observed in both humans and mice, these relationships diverged with respect to PFKP. There are seven SNPs in PFKP that cause coding variation across the eight JDO founder strains, but not even a suggestive association was observed in the region (**Supplementary Figure 11**).

RBCs from donors with high end of storage ATP and low hypoxanthine are characterized by lower hemolysis in vitro and in vivo in autologous healthy and non-autologous critically ill recipients requiring transfusion

End of storage spontaneous hemolysis - one of the two FDA gold standards for the quality of stored blood products - was measured in all 12,753 units from the REDS RBC Omics index donor cohort. Lactate levels at storage day 42 were the top predictor of spontaneous hemolysis (**Figure 4.A**). Likewise, lactate and hypoxanthine levels ranked as the top positive correlates and glucose and ATP among the top 10 negative correlates to the total number of vesicles measured in the REDS RBC Omics blood units at the end of storage (**Figure 4.B**). In a separate cohort of 79 volunteer from the

Donor Iron Deficiency Study (DIDS) in frequent donors,[27] glycolytic metabolites and ATP levels were positively associated to, and hypoxanthine negatively associated to measurements of post-transfusion recovery with ^{51}Cr-labeled RBCs in healthy autologous volunteers – the second and final FDA gold standard to monitor the quality of stored RBCs (**Figure 4.C**). By accessing the "vein-to-vein" database of the REDS RBC Omics program, we first linked blood donor PFKP SNPs to decreases in adjusted hemoglobin increments as a function of allele frequency for rs61835134, the top genetic predictor of ATP levels, in clinical recipients of single irradiated blood unit transfusion (**Figure 4.D**). We then linked ATP and hypoxanthine levels to hemoglobin increments (n=5,816 records available in the database) and bilirubin increments (n=914) upon heterologous transfusion of products from the same donors to critically ill recipients requiring transfusion (**Figure 4.E**). Results indicate a significant positive association between ATP levels and hemoglobin increments in transfusion recipients, with decreased bilirubinemia (lower in vivo hemolysis) in patients receiving unis from donors with high end of storage ATP. Vice versa, opposite results were observed for hypoxanthine, a negative regulator of in vivo hemoglobin increments and elevated bilirubin increments.

DISCUSSION

Herein we report the largest metabolomics study to date of stored RBCs from 13,029 donor volunteers. Results show that end of storage units are metabolically heterogenous, and we identify a cluster of "super"-donors whose RBCs are characterized by elevated glycolysis and ATP levels, and low hypoxanthine levels. This is relevant in that ATP fuels all key processes in mature RBCs, from hemoglobin allostery[28] to metabolism[2,3], from proton pumps to membrane integrity via phosphorylation of structural proteins, from protein stabilization via fueling of transglutaminase 2[29] to cellular mechanics[30], from cytoskeletal actin polymerization[31] to vesiculation[32], from membrane regulation of lipid asymmetry by fueling phosphatidylserine flippases[33] to proteasomal activity to remove damaged proteins.[34] It has long been posited that the ATP-depleted erythrocyte is rapidly removed from the bloodstream[2] via mechanisms of intra- or extravascular hemolysis, such as sequestration in the splenic slit and erythrophagocytosis by resident macrophages in the reticuloendothelial system.[35]

Specifically, in the present study we show that age is a negative contributor to glycolysis, which is also higher in male, non-obese donors compared to pre-menopausal and obese females. After observing ethnicity-based differences in glycolytic rates, as inferred from steady state measurements, we performed mQTL analyses that identified common polymorphisms in PFKP (followed by HK1)

as the main contributors to glycolytic heterogeneity in human blood donor volunteers. While PFK and HK are known rate-limiting enzymes of glycolysis,[1] it had not been documented that the regions coding for these genes are polymorphic in healthy humans. These polymorphisms affect stored RBC metabolism and influence hemolysis in vitro and in vivo, the latter following transfusion into healthy autologous volunteers and non-autologous transfusion into thousands of critically ill patients. Of note, the units tested here were all leukocyte (log4.5) and platelet (log2.5) depleted, making the PFKP hit unlikely to result from an enrichment in platelets in a subset of units.

As a novel finding in this study, we identified an association of polymorphisms in the ectonucleotidase CD38 with the regulation of NAD(P), pentose phosphate pathway and late glycolysis (downstream to glyceraldehyde 3-phosphate) in humans and mice from the J:DO population. While literature suggests a role for CD38 in the regulation of intracellular calcium pools via metabolism of ADP-ribose,[36,37] our results do not show a correlation with calcium levels, as measured via ICP-MS. On the other hand, the positive association with ribose phosphate, metabolites in the non-oxidative phase of the pentose phosphate pathway and late, payoff steps of glycolysis are suggestive of a role for NAD(P) breakdown in stored RBC units as a fuel for ATP synthesis upon phosphorolysis of ADP-ribose and re-entering of the pentose phosphate into glycolysis downstream to glyceraldehyde 3-phosphate.

This finding, here validated in hundreds of genetically diverse mice and over ten thousand humans, may be relevant to other cell types in health and disease. While small scale studies in twins had suggested the heritability of RBC ATP levels,[38] no study to date had addressed the impact of biological and genetic factors influencing energy metabolism in stored erythrocytes. RBC storage is not only a clinically relevant manipulation of the erythrocyte, but also a form of stress (especially, oxidative stress[12]) that at least partially recapitulates other biologically-relevant stressors, such as exercise,[39] exposure to drugs (e.g., chemotherapeutic anthracyclines like doxorubicin[40]), irradiation, or disease (e.g., hemoglobinopathies like sickle cell disease[41] or chronic kidney disease).[3]

By leveraging a massive vein-to-vein database,[42] our findings corroborate real-world clinical data generated from decades of studies suggesting a role for ATP levels in the regulation of RBC hemolytic propensity, especially extravascular hemolysis upon transfusion.[43] Similarly, it had been recently reported that hypoxanthine accumulation as a function of adenosine or AMP deaminase activation during storage is a marker of the metabolic storage lesion that predicts extravascular hemolysis in mice and humans.[44] Here we confirm this laboratory evidence with our autologous PTR studies, and then provide the first clinical data in support to these findings in thousands of transfusion recipients from vein-to-vein database analyses. Since RBC ATP levels are also essential regulators of oxygen kinetics,[4] the present study represents a step forward towards the identification of

advantageous polymorphisms in the determination of "super"-donors not just with respect to decreased hemolytic propensity, but also improved reperfusion performances in vivo in patients requiring transfusion. As the promise of next-generation ex vivo farmed RBCs from hematopoietic or induced-pluripotent stem cell sources materializes,[45] the identification of optimal starting genotypes not just with respect to blood group and rare antigen compatibility, but also with boosted energy metabolism could inform the engineering of ex vivo farmed blood units with optimal storage and resuscitation performances. These considerations may also hold true when considering other relevant civilian or military settings (e.g., athletes, soldiers), for example for the selection of soldiers facing challenges requiring increased oxygen consumption, such as exposure to high altitude hypoxia, strenuous or prolonged physical activities or hemorrhage.

Finally, RBCs and well-curated cohorts of repeated healthy blood donors represent a perfect model to investigate glycolysis at the population levels. Indeed, RBCs are a cell type that exclusively relies on glycolysis to meet its energy demands. Moreover, blood storage is a clinically-relevant stress-intervention that exacerbates the reliance on glycolytic fluxes in the absence of compensatory mechanisms from other organs in vivo. We argue that lessons learned from these studies are translationally relevant for many other research endeavors. Indeed, metabolic reprogramming towards glycolysis – often referred to as Warburg effect in the context of cancer – is perhaps one of the most promising areas of research in the last two decades. Glycolytic reprogramming has been linked as one of the key etiological contributors to many diseases, from cancer[46] to cardiovascular disease,[47] from immunometabolism[48] to aging[49] and age-related neurodegenerative diseases.[50] Our findings suggest that the severity of glycolytic reprogramming in humans - e.g., upon accumulation of mutations to oncogenes/tumor suppressor genes or in response to other factors (e.g., inflammation) - may be mediated by the same biological and genetic factors that regulate glycolysis in human RBCs as described in the present study, making our findings likely relevant to multiple other disciplines beyond transfusion medicine.

References

1 Tanner, L. B. *et al.* Four Key Steps Control Glycolytic Flux in Mammalian Cells. *Cell Syst* **7**, 49-62.e48, doi:10.1016/j.cels.2018.06.003 (2018).

2 van Wijk, R. & van Solinge, W. W. The energy-less red blood cell is lost: erythrocyte enzyme abnormalities of glycolysis. *Blood* **106**, 4034-4042, doi:10.1182/blood-2005-04-1622 (2005).

3 D'Alessandro, A. *et al.* Red Blood Cell Metabolism In Vivo and In Vitro. *Metabolites* **13**, doi:10.3390/metabo13070793 (2023).

4 Donovan, K. *et al.* Stored blood has compromised oxygen unloading kinetics that can be normalized with rejuvenation and predicted from corpuscular side-scatter. *Haematologica*, doi:10.3324/haematol.2021.279296 (2021).

5 Churchill, G. A., Gatti, D. M., Munger, S. C. & Svenson, K. L. The Diversity Outbred mouse population. *Mamm Genome* **23**, 713-718, doi:10.1007/s00335-012-9414-2 (2012).

6 Wiback, S. J. & Palsson, B. O. Extreme pathway analysis of human red blood cell metabolism. *Biophys J* **83**, 808-818, doi:10.1016/s0006-3495(02)75210-7 (2002).

7 Grüning, N.-M. & Ralser, M. Glycolysis: How a 300yr long research journey that started with the desire to improve alcoholic beverages kept revolutionizing biochemistry. *Current Opinion in Systems Biology* **28**, 100380, doi:https://doi.org/10.1016/j.coisb.2021.100380 (2021).

8 Meyerhof, O. & Junowicz-Kocholaty, R. THE EQUILIBRIA OF ISOMERASE AND ALDOLASE, AND THE PROBLEM OF THE PHOSPHORYLATION OF GLYCERALDEHYDE PHOSPHATE. *Journal of Biological Chemistry* **149**, 71-92, doi:https://doi.org/10.1016/S0021-9258(18)72218-7 (1943).

9 Bianconi, E. *et al.* An estimation of the number of cells in the human body. *Ann Hum Biol* **40**, 463-471, doi:10.3109/03014460.2013.807878 (2013).

10 Nemkov, T., Reisz, J. A., Xia, Y., Zimring, J. C. & D'Alessandro, A. Red blood cells as an organ? How deep omics characterization of the most abundant cell in the human body highlights other systemic metabolic functions beyond oxygen transport. *Expert review of proteomics* **15**, 855-864, doi:10.1080/14789450.2018.1531710 (2018).

11 Kaestner, L. & Minetti, G. The potential of erythrocytes as cellular aging models. *Cell Death Differ* **24**, 1475-1477, doi:10.1038/cdd.2017.100 (2017).

12 Yoshida, T., Prudent, M. & D'Alessandro, A. Red blood cell storage lesion: causes and potential clinical consequences. *Blood Transfus* **17**, 27-52, doi:10.2450/2019.0217-18 (2019).

13 Isiksacan, Z. *et al.* Assessment of stored red blood cells through lab-on-a-chip technologies for precision transfusion medicine. *Proceedings of the National Academy of Sciences* **120**, e2115616120, doi:doi:10.1073/pnas.2115616120 (2023).

14 Page, G. P. *et al.* Multiple-ancestry genome-wide association study identifies 27 loci associated with measures of hemolysis following blood storage. *J Clin Invest* **131**, doi:10.1172/jci146077 (2021).

15 Roubinian, N. H. *et al.* Donor genetic and nongenetic factors affecting red blood cell transfusion effectiveness. *JCI Insight* **7**, doi:10.1172/jci.insight.152598 (2022).

16 Surendran, P. *et al.* Rare and common genetic determinants of metabolic individuality and their effects on human health. *Nature Medicine* **28**, 2321-2332, doi:10.1038/s41591-022-02046-0 (2022).

17 Childs, B. Sir Archibald Garrod's Conception of Chemical Individuality: A Modern Appreciation. *New England Journal of Medicine* **282**, 71-77, doi:10.1056/nejm197001082820205 (1970).

18 Draisma, H. H. M. *et al.* Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. *Nature Communications* **6**, 7208, doi:10.1038/ncomms8208 (2015).

19 Appiah, D., Nwabuo, C. C., Ebong, I. A., Wellons, M. F. & Winters, S. J. Trends in Age at Natural Menopause and Reproductive Life Span Among US Women, 1959-2018. *JAMA* **325**, 1328-1330, doi:10.1001/jama.2021.0278 (2021).

20 Sanghvi, V. R. *et al.* The Oncogenic Action of NRF2 Depends on De-glycation by Fructosamine-3-Kinase. *Cell* **178**, 807-819.e821, doi:10.1016/j.cell.2019.07.031 (2019).

21 Mosca, L. *et al.* Genetic variability of the fructosamine 3-kinase gene in diabetic patients. *Clin Chem Lab Med* **49**, 803-808, doi:10.1515/cclm.2011.133 (2011).

22 Ceroni, F. *et al*. A deletion involving CD38 and BST1 results in a fusion transcript in a patient with autism and asthma. *Autism Res* **7**, 254-263, doi:10.1002/aur.1365 (2014).
23 Quarona, V. *et al.* CD38 and CD157: a long journey from activation markers to multifunctional molecules. *Cytometry B Clin Cytom* **84**, 207-217, doi:10.1002/cyto.b.21092 (2013).
24 El-Ghamrawy, M. *et al.* Association between BCL11A, HSB1L-MYB, and XmnI γG-158 (C/T) gene polymorphism and hemoglobin F level in Egyptian sickle cell disease patients. *Ann Hematol* **99**, 2279-2288, doi:10.1007/s00277-020-04187-z (2020).
25 Babbs, C. *et al.* ATR-16 syndrome: mechanisms linking monosomy to phenotype. *J Med Genet* **57**, 414-421, doi:10.1136/jmedgenet-2019-106528 (2020).
26 Howie, H. L. *et al.* Differences in Steap3 expression are a mechanism of genetic variation of RBC storage and oxidative damage in mice. *Blood Advances* **3**, 2272-2285, doi:10.1182/bloodadvances.2019000605 (2019).
27 Hod, E. A. *et al.* A randomized trial of blood donor iron repletion on red cell quality for transfusion and donor cognition and well-being. *Blood* **140**, 2730-2739, doi:10.1182/blood.2022017288 (2022).
28 Ahmed, M. H., Ghatge, M. S. & Safo, M. K. Hemoglobin: Structure, Function and Allostery. *Subcell Biochem* **94**, 345-382, doi:10.1007/978-3-030-41769-7_14 (2020).
29 Xu, P. *et al.* Erythrocyte transglutaminase-2 combats hypoxia and chronic kidney disease by promoting oxygen delivery and carnitine homeostasis. *Cell Metab* **34**, 299-316.e296, doi:10.1016/j.cmet.2021.12.019 (2022).
30 Betz, T., Lenz, M., Joanny, J.-F. & Sykes, C. ATP-dependent mechanics of red blood cells. *Proceedings of the National Academy of Sciences* **106**, 15320-15325, doi:doi:10.1073/pnas.0904614106 (2009).
31 Gokhin, D. S. & Fowler, V. M. Feisty filaments: actin dynamics in the red blood cell membrane skeleton. *Curr Opin Hematol* **23**, 206-214, doi:10.1097/moh.0000000000000227 (2016).
32 Wagner, G. M., Chiu, D. T., Yee, M. C. & Lubin, B. H. Red cell vesiculation--a common membrane physiologic event. *J Lab Clin Med* **108**, 315-324 (1986).
33 Arashiki, N. *et al.* ATP11C is a major flippase in human erythrocytes and its defect causes congenital hemolytic anemia. *Haematologica* **101**, 559-565, doi:10.3324/haematol.2016.142273 (2016).
34 Tzounakas, V. L. *et al.* Red cell proteasome modulation by storage, redox metabolism and transfusion. *Blood Transfus* **20**, 27-39, doi:10.2450/2020.0179-20 (2022).
35 D'Alessandro, A. *et al.* Red Blood Cell Metabolism In Vivo and In Vitro. *Metabolites* **13**, 793 (2023).
36 Lee, H. C., Deng, Q. W. & Zhao, Y. J. The calcium signaling enzyme CD38 - a paradigm for membrane topology defining distinct protein functions. *Cell Calcium* **101**, 102514, doi:https://doi.org/10.1016/j.ceca.2021.102514 (2022).
37 De Flora, A., Zocchi, E., Guida, L., Franco, L. & Bruzzone, S. Autocrine and paracrine calcium signaling by the CD38/NAD+/cyclic ADP-ribose system. *Ann N Y Acad Sci* **1028**, 176-191, doi:10.1196/annals.1322.021 (2004).
38 van 't Erve, T. J. *et al.* The heritability of metabolite concentrations in stored human red blood cells. *Transfusion* **54**, 2055-2063, doi:10.1111/trf.12605 (2014).
39 Nemkov, T. *et al.* Metabolic Signatures of Performance in Elite World Tour Professional Male Cyclists. *Sports Med*, 1-15, doi:10.1007/s40279-023-01846-9 (2023).
40 Bissinger, R. *et al.* Proteinuric chronic kidney disease is associated with altered red blood cell lifespan, deformability and metabolism. *Kidney Int* **100**, 1227-1239, doi:10.1016/j.kint.2021.08.024 (2021).
41 D'Alessandro, A. *et al.* In vivo evaluation of the effect of sickle cell hemoglobin S, C and therapeutic transfusion on erythrocyte metabolism and cardiorenal dysfunction. *Am J Hematol*, DOI: 10.1002/ajh.26923, doi:10.1002/ajh.26923 (2023).
42 Roubinian, N. H. *et al.* Effect of donor, component, and recipient characteristics on hemoglobin increments following red blood cell transfusion. *Blood* **134**, 1003-1013, doi:10.1182/blood.2019000773 (2019).
43 Heaton, W. A., Keegan, T., Holme, S. & Momoda, G. Evaluation of 99mtechnetium/51chromium post-transfusion recovery of red cells stored in saline, adenine, glucose, mannitol for 42 days. *Vox Sang* **57**, 37-42, doi:10.1111/j.1423-0410.1989.tb04981.x (1989).
44 Nemkov, T. *et al.* Hypoxia modulates the purine salvage pathway and decreases red blood cell and supernatant levels of hypoxanthine during refrigerated storage. *Haematologica* **103**, 361-372, doi:10.3324/haematol.2017.178608 (2018).

45 Trakarnsanga, K. *et al.* An immortalized adult human erythroid line facilitates sustainable and scalable generation of functional red cells. *Nat Commun* **8**, 14750, doi:10.1038/ncomms14750 (2017).
46 Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029-1033, doi:10.1126/science.1160809 (2009).
47 D'Alessandro, A. *et al.* Hallmarks of Pulmonary Hypertension: Mesenchymal and Inflammatory Cell Metabolic Reprogramming. *Antioxid Redox Signal* **28**, 230-250, doi:10.1089/ars.2017.7217 (2018).
48 Pålsson-McDermott, E. M. & O'Neill, L. A. J. Targeting immunometabolism as an anti-inflammatory strategy. *Cell Research* **30**, 300-314, doi:10.1038/s41422-020-0291-z (2020).
49 Loeb, L. A., Wallace, D. C. & Martin, G. M. The mitochondrial theory of aging and its relationship to reactive oxygen species damage and somatic mtDNA mutations. *Proceedings of the National Academy of Sciences* **102**, 18769-18770, doi:doi:10.1073/pnas.0509776102 (2005).
50 Traxler, L. *et al.* Warburg-like metabolic transformation underlies neuronal degeneration in sporadic Alzheimer's disease. *Cell Metab* **34**, 1248-1263.e1246, doi:10.1016/j.cmet.2022.07.014 (2022).

METHODS

Donor recruitment in the REDS RBC Omics study:

Index donors A total of 13,758 donors were enrolled in the Recipient Epidemiology and Donor evaluation Study (REDS) RBC Omics at four different blood centers across the United States (https://biolincc.nhlbi.nih.gov/studies/reds_iii/). Of these, 97% (13,403) provided informed consent and 13,029 were available for metabolomics analyses in this study – henceforth referred to as "index donors". A subset of these donors were evaluable for hemolysis parameters, including spontaneous (n=12,753) in ~42-day stored leukocyte-filtered packed RBCs derived from whole blood donations from this cohort.[51] Methods for the determination of FDA-standard spontaneous (storage) hemolysis test, osmotic hemolysis (pink test) and oxidative hemolysis upon challenge with AAPH have been extensively described elsewhere.[52]

Storage hemolysis In details, after storage under routine blood bank conditions (1-6°C) for 39 to 42 days, the contents of each unit was transferred into a 15-mL conical tube from which 2 aliquots (1 mL each) were processed for the hemolytic assays. One aliquot was used for the quantification of spontaneous storage hemolysis and the other for the stress-induced hemolysis assays. Percent end-of-storage hemolysis was determined according to the following equation:

$$\textit{Storage hemolysis (\%)} = (100\text{-}HCT)\ Hb_{supernatant} : Hb_{total}$$

Sample hematocrit (HCT) was determined by collecting blood samples into capillary tubes, which were centrifuged in a micro-HCT centrifuge (LW Scientific, Lawrenceville, GA). $Hb_{supernatant}$ refers to the levels of free hemoglobin obtained after centrifugation (1500g, 10 minutes, 18°C) measured in the supernatant. Hb_{total} refers to the total amount of sample hemoglobin before centrifugation. In the entire study, hemoglobin concentrations (micromolar) were determined by Drabkin's method.[53]

For the evaluation of stress-induced hemolysis, stored RBCs were washed (1500g, 10 minutes, 18°C) 3 times with phosphate-buffered saline (PBS) to remove plasma and additive solution, and immediately subjected to osmotic or oxidative stress assays.

RBC osmotic hemolysis

Osmotic hemolysis was determined by a modified pink test assay previously used for diagnosis of genetic mutations that affect RBC osmotic fragility (eg, sickle cell disease, thalassemia, and spherocytosis).[54,55] Washed RBCs were incubated under static conditions (4 hours at 22°C) in pink test buffer (a hypotonic Bis-Tris buffer containing 25 mmol/L sodium chloride, 70 mmol/L 2,2-Bis(hydroxymethyl)-2,2′,2″nitrilotriethanol (Bis-Tris) buffer, and 135 mmol/L glycerol; pH 6.6) at a final concentration of 1.6% ± 0.2% after which samples were centrifuged (1500g, 10 minutes, 18°C), and percent osmotic hemolysis was determined:

$$\textit{Osmotic hemolysis (\%)} = (Hb_{Osmotic} - Hb_{total})\ x\ 100$$

for which Hbosmotic corresponds to supernatant cell-free hemoglobin of pink test–treated RBCs, and Hb_{total} refers to the total amount of hemoglobin in each sample.

RBC oxidative hemolysis

Testing for donor differences in RBC susceptibility to oxidative hemolysis was performed as described by Kanias et al.,[52] by incubating RBCs in the presence of 2,2′-azobis-2-methyl-propanimidamide dihydrochloride (AAPH; 150 mmoL). Thermal (37°C) decomposition of AAPH generates peroxyl radicals leading to lipid peroxidation–mediated hemolysis.24 Washed RBCs were suspended with PBS to a final concentration of 3.5% ± 0.5%. Aliquots (0.21 mL) were transferred

into microplates to which 0.09 mL of AAPH (0.5 M) or PBS was added. The plates were incubated (37°C) under static conditions for 1.5 hours, after which the plates were centrifuged (1500g, 10 minutes, 18°C), and AAPH-induced oxidative hemolysis was determined by using the following formula:

$$\textit{Oxidative hemolysis (\%)} = (Hb_{AAPH} - Hb_{control}) : Hb_{total} \times 100$$

where Hb_{AAPH} corresponds to supernatant cell-free hemoglobin of AAPH-treated RBCs, $Hb_{control}$ corresponds to supernatant cell-free hemoglobin from untreated RBCs, and Hb_{total} refers to the total amount of hemoglobin in each sample.

Recalled donors: A total of 643 donors scoring in the 5th and 95th percentile for hemolysis parameters at the index phase of the study were invited to donate a second unit of pRBCs, a cohort henceforth referred to as "recalled donors". These units were assayed at storage days 10, 23 and 42 for hemolytic parameters and mass spectrometry-based high-throughput metabolomics,[56] proteomics,[57] lipidomics[58] and ICP-MS analyses[59]. Under the aegis of the REDS-IV-P project,[60] a total of 1,929 samples (n=643, storage day 10, 23 and 42) were processed with this multi-omics workflow. Extracellular vesicles were measured by flow-cytometry.[61]

Donor iron deficiency study (DIDS) A total of 79 repeat blood donors were enrolled as part of the DIDS study for baseline and post-iron repletion before determination of autologous post-transfusion recovery (PTR) studies of day 42 stored units with ^{51}Cr, as described.[27]

High-throughput metabolomics Metabolomics extraction and analyses in 96 well-plate format were performed as described.[62,63] RBC samples were transferred on ice on 96 well plate and frozen at -80 ºC at Vitalant San Francisco prior to shipment in dry ice to the University of Colorado Anschutz Medical Campus. Plates were thawed on ice then a 10 uL aliquot was transferred with a multi-channel pipettor to 96-well extraction plates. A volume of 90 uL of ice cold 5:3:2 MeOH:MeCN:water (*v/v/v*) was added to each well, with an electronically-assisted cycle of sample mixing repeated three times. Extracts were transferred to 0.2 µm filter plates (Biotage) and insoluble material was removed under positive pressure using nitrogen applied via a 96-well plate manifold. Filtered extracts were transferred to an ultra-high-pressure liquid chromatography (UHPLC-MS — Vanquish) equipped with a plate charger. A blank containing a mix of standards detailed before[64] and a quality control sample (the same across all plates) were injected 2 or 5 times each per plate, respectively, and used to monitor instrument performance throughout the analysis. Metabolites were resolved on a Phenomenex Kinetex C18 column (2.1 x 30 mm, 1.7 um) at 45 °C using a 1-minute ballistic gradient method in positive and negative ion modes (separate runs) over the scan range 65-975 m/z exactly as previously described.[62] The UHPLC was coupled online to a Q Exactive mass spectrometer (Thermo Fisher). The Q Exactive MS was operated in negative ion mode, scanning in Full MS mode (2 µscans) from 90 to 900 m/z at 70,000 resolution, with 4 kV spray voltage, 45 sheath gas, 15 auxiliary gas. Following data acquisition, .raw files were converted to .mzXML using RawConverter then metabolites assigned and peaks integrated using ElMaven (Elucidata) in conjunction with an in-house standard library.[65]

mQTL analysis in REDS RBC Omics The workflow for the mQTL analysis of ATP, hypoxanthine and glycolytic metabolites is consistent with previously described methods from our pilot mQTL study on 250 recalled donors.[66] Details of the genotyping and imputation of the RBC

Omics study participants have been previously described by Page, et al.[14] Briefly, genotyping was performed using a Transfusion Medicine microarray[67] consisting of 879,000 single nucleotide polymorphisms (SNPs); the data are available in dbGAP accession number phs001955.v1.p1. Imputation was performed using 811,782 SNPs that passed quality control. After phasing using Shape-IT,[68] imputation was performed using Impute2[69] with the 1000 Genomes Project phase 3[69] all-ancestry reference haplotypes. We used the R package SNPRelate[70] to calculate principal components (PCs) of ancestry. We performed association analyses for kynurenine using an additive SNP model in the R package ProbABEL[71] and 13,091 study participants who had both metabolomics data and imputation data on serial samples from stored RBC components that passed respective quality control procedures. We adjusted for sex, age (continuous), frequency of blood donation in the last two years (continuous), blood donor center, and ten ancestry PCs. Statistical significance was determined using a p-value threshold of $5x10^{-8}$. We only considered variants with a minimum minor allele frequency of 1% and a minimum imputation quality score of 0.80. The OASIS: Omics Analysis, Search & Information a TOPMED funded resources[72], was used to annotate the top SNPs. OASIS annotation includes information on position, chromosome, allele frequencies, closest gene, type of variant, position relative to closest gene model, if predicted to functionally consequential, tissues specific gene expression, and other information.

Mouse RBC storage and post-transfusion recovery Mouse storage and PTR studies were performed as previously described,[73] leveraging a population of 350 Diversity Outbred mice obtained from the Jackson Laboratory (J:DO). J:DO mice were derived from extensive outbreeding of eight inbred founder strains that represent genetically distinct lineages of the house mouse.[5] All animal procedures were approved by the University of Virginia IACUC (protocol no. 4269). All animals were genotyped at 143,259 SNPs (137,192 loci after filtering) using the GigaMUGA array.[74]

mQTL analysis in JAX DO mice The mQTL workflow in JAX DO mice followed previously defined conventions.[75] Briefly, metabolite values of zero were converted to a missing value and only metabolites with 100 non-missing observations or more were included for further analyses. Each metabolite was transformed to normal quantiles for mQTL analysis to reduce the influence of outlying values. The initial mQTL mapping was based on founder strain haplotypes imputed at SNPs (137,192 loci), allowing for the additive genetic effects at loci to be more richly characterized as eight founder allele effects. Sex (148 females and 202 males) and sub-cohorts (5 groups ranging from 27 to 94 mice) were adjusted for as covariates. A lenient significance threshold of LOD score > 6 was used for calling mQTL to allow for the detection of mQTL hot spots. For reference, a LOD score > 8 approximately represents a genome-wide adjusted p-value < 0.05.[76] Fine mapping at detected mQTL was performed by imputing variants based on the genotypes of the founder strains (GRCm39). We used the same mQTL workflow for fresh and stored RBCs. All analyses were performed using the R package qtl2.[75]

Determination of hemoglobin and bilirubin increment via the vein-to-vein database Association of metabolite levels with hemoglobin and bilirubin increment was performed by interrogating the vein-to-vein database, as described in Roubinian et al.[15,42] Bucketing of metabolite levels in quartiles and tertiles for hemoglobin and bilirubin increments was informed by the significantly smaller record of entries for the latter parameter in the database.

Vein-to-vein database: General Study Design We conducted a retrospective cohort study using electronic health records from the National Heart Lung and Blood Institute (NHLBI) Recipient Epidemiology and Donor Evaluation Study-III (REDS-III) program available as public use data through BioLINCC.[77,78] The database includes blood donor, component manufacturing, and patient data collected at 12 academic and community hospitals from four geographically diverse regions in the United States (Connecticut, Pennsylvania, Wisconsin, and California) for the 4-year period from January 1, 2013 to December 31, 2016. Genotype and metabolomic data from the subset of blood donors who participated in the REDS-III RBC-Omics study[79] was linked to the dataset using unique donor identifiers.

Study Population and Definitions Data regarding blood donor demographics (e.g., age, sex, body mass index, ABO/Rh), collection date and methods (e.g., whole blood or apheresis), as well as component manufacturing characteristics (e.g., timing of leukoreduction, additive solution, irradiation status) were extracted for each RBC unit collected. Available donor genetic polymorphism and metabolomic data was linked to issued RBC units using random donor identification numbers. Among transfusion recipients, we included all adult patients who received a single RBC unit during one or more transfusion episodes between January 1, 2013 and December 30, 2016. Recipient details included age, sex, body mass index (BMI), along with storage age, and blood product issue date and time. We collected hemoglobin and bilirubin levels measured by the corresponding clinical laboratory prior to and following each RBC transfusion event.

Transfusion Exposures and Outcome Measures All single RBC unit transfusion episodes linked to genetic polymorphism and metabolomic data were included in this analysis. A RBC unit transfusion episode was defined as any single RBC transfusion from a single donor with both informative pre- and post-transfusion laboratory measures and without any other RBC units transfused in the following 24-hour time period. The outcome measures of interest were change in hemoglobin (ΔHb; g/dL) and change in total bilirubin (ΔBi; mg/dL) following a single RBC unit transfusion episode. These outcomes were defined as the difference between the post-transfusion and pre-transfusion levels. Pre-transfusion thresholds for these measures were chosen to limit patient confounding (e.g., underlying hepatic disease). For pre-transfusion hemoglobin, the value used was the most proximal hemoglobin measurements prior to RBC transfusion, but at most 24 hours prior to transfusion. Furthermore, we excluded transfusion episodes where the pre-transfusion hemoglobin was greater than 9.5 g/dL, and the hemoglobin increment may be confounded by hemorrhage events. For post-transfusion hemoglobin, the laboratory measure nearest to 24-hours post-transfusion, but between 12- and 36-hours following transfusion was used. For pre-transfusion bilirubin, the most proximal total bilirubin measurement prior to transfusion, but at most 7-days prior to transfusion was used. Furthermore, the pre-transfusion bilirubin had to be below 1.5 mg/dL for the transfusion to be considered to exclude patients with pre-existing liver disease. For post-transfusion bilirubin, the most proximal laboratory measure to the transfusion, but at most 24-hours post-transfusion was used.

Statistical Methods We assessed the univariable association of all *a priori* selected donor, manufacturing, and recipient covariates with the outcomes using linear regression. Multivariable linear regression assessed associations between alleles of donor single nucleotide polymorphisms and changes in hemoglobin and bilirubin levels post-transfusion hemoglobin. Two-sided p-values less

than 0.05 were considered to be statistically significant. Analyses were performed using Stata Version 14.1, StataCorp, College Station, TX.

Data analysis Statistical analyses – including hierarchical clustering analysis (HCA), linear discriminant analysis (LDA), uniform Manifold Approximation and Projection (uMAP), correlation analyses and Lasso regression were performed using both MetaboAnalyst 5.0[80] and in-house developed code in R (4.2.3 2023-03-15).

REFERENCES (METHODS only)

51 D'Alessandro, A. *et al.* Heterogeneity of blood processing and storage additives in different centers impacts stored red blood cell metabolism as much as storage time: lessons from REDS-III-Omics. *Transfusion* **59**, 89-100, doi:10.1111/trf.14979 (2019).

52 Kanias, T. *et al.* Ethnicity, sex, and age are determinants of red blood cell storage and stress hemolysis: results of the REDS-III RBC-Omics study. *Blood Adv* **1**, 1132-1141, doi:10.1182/bloodadvances.2017004820 (2017).

53 Zwart, A. *et al.* Recommendations for reference method for haemoglobinometry in human blood (ICSH standard 1995) and specifications for international haemiglobinocyanide standard (4th edition). *J Clin Pathol* **49**, 271-274, doi:10.1136/jcp.49.4.271 (1996).

54 Judkiewicz, L., Bugala, I. & Bartosz, G. 'Pink test' and osmotic fragility test for the diagnosis of hereditary spherocytosis: another view. *Eur J Haematol* **42**, 217, doi:10.1111/j.1600-0609.1989.tb01216.x (1989).

55 Vettore, L. *et al.* A new test for the laboratory diagnosis of spherocytosis. *Acta Haematol* **72**, 258-263, doi:10.1159/000206398 (1984).

56 Nemkov, T., Yoshida, T., Nikulina, M. & D'Alessandro, A. High-Throughput Metabolomics Platform for the Rapid Data-Driven Development of Novel Additive Solutions for Blood Storage. *Frontiers in Physiology* **13**, doi:10.3389/fphys.2022.833242 (2022).

57 Thomas, T. *et al.* Evidence for structural protein damage and membrane lipid remodeling in red blood cells from COVID-19 patients. *medRxiv*, doi:10.1101/2020.06.29.20142703 (2020).

58 Reisz, J. A., Zheng, C., D'Alessandro, A. & Nemkov, T. Untargeted and Semi-targeted Lipid Analysis of Biological Samples Using Mass Spectrometry-Based Metabolomics. *Methods Mol Biol* **1978**, 121-135, doi:10.1007/978-1-4939-9236-2_8 (2019).

59 Stephenson, D., Nemkov, T., Qadri, S. M., Sheffield, W. P. & D'Alessandro, A. Inductively-Coupled Plasma Mass Spectrometry-Novel Insights From an Old Technology Into Stressed Red Blood Cell Physiology. *Front Physiol* **13**, 828087, doi:10.3389/fphys.2022.828087 (2022).

60 Josephson, C. D. *et al.* The Recipient Epidemiology and Donor Evaluation Study-IV-Pediatric (REDS-IV-P): A research program striving to improve blood donor safety and optimize transfusion outcomes across the lifespan. *Transfusion* **62**, 982-999, doi:10.1111/trf.16869 (2022).

61 Bakkour, S. *et al.* Manufacturing method affects mitochondrial DNA release and extracellular vesicle composition in stored red blood cells. *Vox Sanguinis* **111**, 22-32, doi:https://doi.org/10.1111/vox.12390 (2016).

62 Nemkov, T., Yoshida, T., Nikulina, M. & D'Alessandro, A. High-Throughput Metabolomics Platform for the Rapid Data-Driven Development of Novel Additive Solutions for Blood Storage. *Front Physiol* **13**, 833242, doi:10.3389/fphys.2022.833242 (2022).

63 Nemkov, T., Reisz, J. A., Gehrke, S., Hansen, K. C. & D'Alessandro, A. High-Throughput Metabolomics: Isocratic and Gradient Mass Spectrometry-Based Methods. *Methods Mol Biol* **1978**, 13-26, doi:10.1007/978-1-4939-9236-2_2 (2019).

64 Stefanoni, D. *et al.* Red blood cell metabolism in Rhesus macaques and humans: comparative biology of blood storage. *Haematologica* **105**, 2174-2186, doi:10.3324/haematol.2019.229930 (2020).

65 Nemkov, T., Hansen, K. C. & D'Alessandro, A. A three-minute method for high-throughput quantitative metabolomics and quantitative tracing experiments of central carbon and nitrogen pathways. *Rapid Commun Mass Spectrom* **31**, 663-673, doi:10.1002/rcm.7834 (2017).
66 Moore, A. *et al.* Genome-wide metabolite quantitative trait loci analysis (mQTL) in red blood cells from volunteer blood donors. *J Biol Chem* **298**, 102706, doi:10.1016/j.jbc.2022.102706 (2022).
67 Guo, Y. *et al.* Development and evaluation of a transfusion medicine genome wide genotyping array. *Transfusion* **59**, 101-111, doi:10.1111/trf.15012 (2019).
68 Delaneau, O., Coulonges, C. & Zagury, J.-F. Shape-IT: new rapid and accurate algorithm for haplotype inference. *BMC Bioinformatics* **9**, 540, doi:10.1186/1471-2105-9-540 (2008).
69 Howie, B., Marchini, J. & Stephens, M. Genotype Imputation with Thousands of Genomes. *G3 Genes|Genomes|Genetics* **1**, 457-470, doi:10.1534/g3.111.001198 (2011).
70 Zheng, X. *et al.* A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326-3328, doi:10.1093/bioinformatics/bts606 (2012).
71 Aulchenko, Y. S., Struchalin, M. V. & van Duijn, C. M. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* **11**, 134, doi:10.1186/1471-2105-11-134 (2010).
72 Perry, J. A., Gaynor, B. J., Mitchell, B. D. & O'Connell, J. R. An Omics Analysis Search and Information System (OASIS) for Enabling Biological Discovery in the Old Order Amish. *bioRxiv*, 2021.2005.2002.442370, doi:10.1101/2021.05.02.442370 (2021).
73 Howie, H. L. *et al.* Differences in Steap3 expression are a mechanism of genetic variation of RBC storage and oxidative damage in mice. *Blood Adv* **3**, 2272-2285, doi:10.1182/bloodadvances.2019000605 (2019).
74 Morgan, A. P. & Welsh, C. E. Informatics resources for the Collaborative Cross and related mouse populations. *Mamm Genome* **26**, 521-539, doi:10.1007/s00335-015-9581-z (2015).
75 Broman, K. W. *et al.* R/qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations. *Genetics* **211**, 495-502, doi:10.1534/genetics.118.301595 (2019).
76 Keele, G. R. Which mouse multiparental population is right for your study? The Collaborative Cross inbred strains, their F1 hybrids, or the Diversity Outbred population. *G3 (Bethesda)* **13**, doi:10.1093/g3journal/jkad027 (2023).
77 https://biolincc.nhlbi.nih.gov/studies/reds_iii/.
78 Karafin, M. S. *et al.* Demographic and epidemiologic characterization of transfusion recipients from four US regions: evidence from the REDS-III recipient database. *Transfusion* **57**, 2903-2913, doi:10.1111/trf.14370 (2017).
79 Endres-Dighe, S. M. *et al.* Blood, sweat, and tears: Red Blood Cell-Omics study objectives, design, and recruitment activities. *Transfusion* **59**, 46-56, doi:10.1111/trf.14971 (2019).
80 Pang, Z. *et al.* MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Research* **49**, W388-W396, doi:10.1093/nar/gkab382 (2021).

Acknowledgements AD was supported by funds by the National Heart, Lung, and Blood Institute (R21HL150032, R01HL146442, R01HL149714, R01HL148151, R01HL161004). The authors acknowledge support from the REDS RBC Omics and REDS-IV-P CTLS programs, sponsored by the National Heart, Lung, and Blood Institute contract 75N2019D00033, and from the NHLBI Recipient Epidemiology and Donor Evaluation Study-III (REDS-III) RBC Omics project, which was supported by NHLBI contracts HHSN2682011-00001I, -00002I, -00003I, -00004I, -00005I, -00006I, -00007I, -00008I, and -00009I. The autologous post-transfusion recovery studies were funded by NHLBI R01HL133049. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' contributions TN, DS, AD performed metabolomics analyses. TN, DS, AD performed method development, quality control and metabolomics data analysis. EE, AM, GPP performed human mQTL analyses; GRK, GAC performed mQTL analyses for the Diversity Outbred mice; AH, JCZ performed the Diversity Outbred mouse studies; MD, KCH performed proteomics studies; SLS, EAH performed post-transfusion recovery studies (DIDS cohort); CE, AK, XD, EE, GRK, GAC, NR, AD performed statistical analyses; NR performed linkage analyses to the vein-to-vein database; MS, PJN, SK, MPB performed and supervised the REDS RBC Omics study; TN, EE, GRK, AK, CE, AD prepared figure panels; AD wrote the first version of the manuscript, that was critically reviewed and approved by all authors.

Competing Interests The authors declare that AD, KCH, TN are founders of Omix Technologies Inc and Altis Biosciences LLC. AD, SLS and TN are Scientific Advisory Board (SAB) members for Hemanext Inc. AD is SAB member for Macopharma Inc. SLS is SAB member for Alcor, Inc and consultant for Tioma, Inc and Team Conveyer Intellectual Properties. SLS is also executive director for Worldwide Initiative for Rh Disease Eradication (WIRhE) and CEO for Ferrous Wheel Consultants, LLC. JCZ is a founder of Svalinn Therapeutics. All the other authors have no conflicts to disclose in relation to this study.

FIGURES

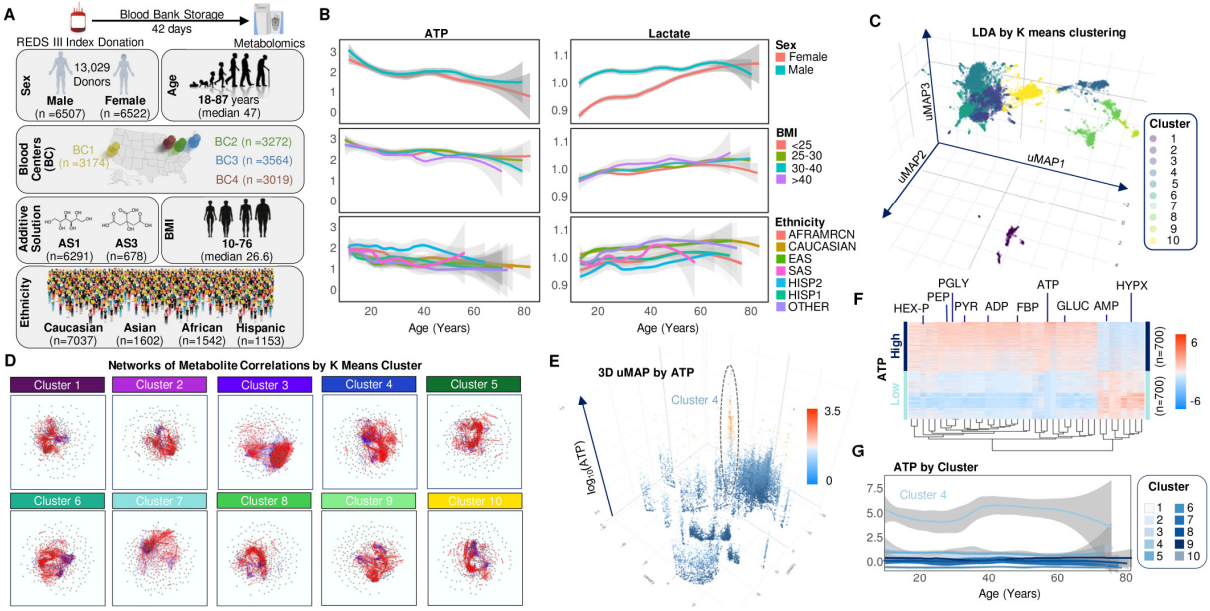


Figure 1 – Metabolomics of 13,029 end of storage packed RBCs from the REDS RBC Omics Index cohort Leukocyte-filtered RBCs were donated by 13,029 volunteers across four different blood centers in the United States and stored for 42 days (end of the shelf-life) prior to testing via high-throughput metabolomics. In **A,** an overview of the study design and donor demographics. In **B,** line plots of representative metabolites ATP and lactate as a function of donor age, sex, body mass index (BMI) and ethnicity. In **C,** linear discriminant analysis color-coded as a function of K mean clustering analysis. In **D**, network view of metabolite-metabolite correlations (Spearman) as a function of the 10 K means clusters. Each metabolite is a node and a red and blue edge indicates positive or negative correlations higher than 0.5 in module. In **E**, Uniform Manifold Approximation and Projection (uMAP) of the metabolomics data from all 13,029 index donors color-coded by ATP levels. In **F**, heat map of the top 50 significant metabolites that discriminate the donors with the lowest and highest ATP levels (5th and 95th percentiles; n=700). In **G**, line plots show that higher ATP levels in cluster 4 are not explained by donor age alone.

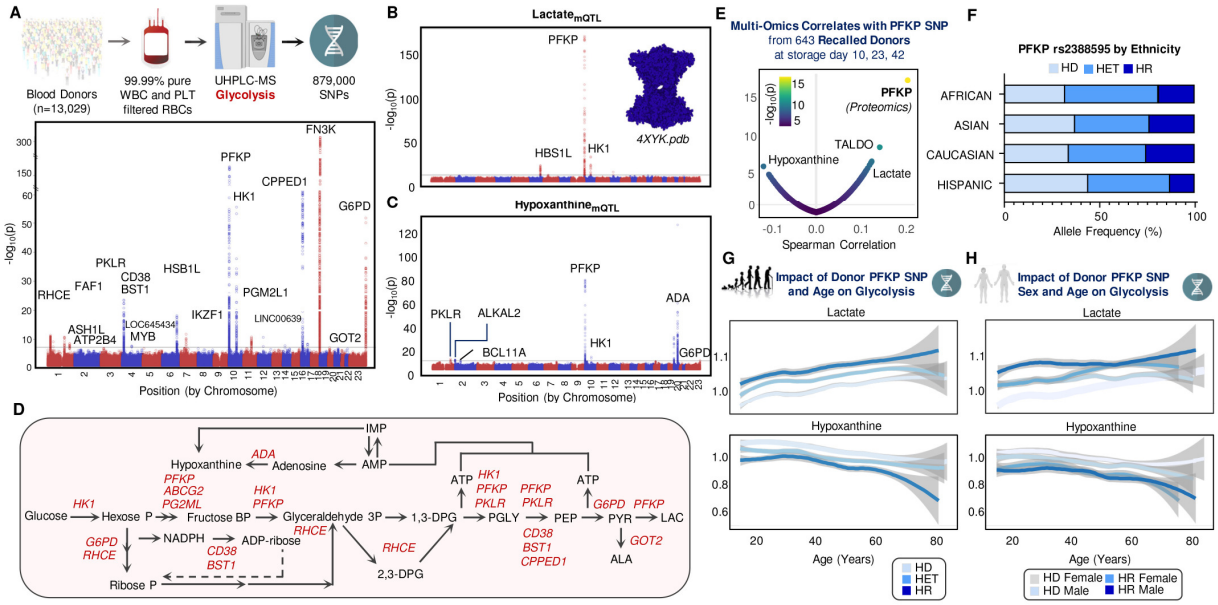


Figure 2 – Genetic underpinnings of glycolytic heterogeneity in end of storage units from 13,029 REDS RBC Omics Index donors. Along with metabolomics analyses, as part of the REDS RBC Omics study we tested donors for ~879,000 single nucleotide polymorphisms (SNPs – **A**). A summary Manhattan plot of metabolite Quantitative Trait Loci (mQTL) is shown in **A** for all glycolytic metabolites merged, or just for lactate (**B**) and hypoxanthine (**C**). In **D**, a summary overview of the gene-metabolite associations identified by the mQTL analysisIn **E**, multi-omics analyses of recalled donor units (n=643, storage day 10, 23, 42) confirmed a strong positive correlation between PFKP SNP rs2388595 alleles and PFKP levels by proteomics, and positive and negative associations with lactate and hypoxanthine. . In **F**, ethnic breakdown of allele frequency for the top PFKP SNPs (rs2388595). In **G-H**, combined effects of PFKP SNP with donor age and sex for lactate and hypoxanthine. HD: homozygous dominant; HET: heterozygous; HR: homozygous recessive.

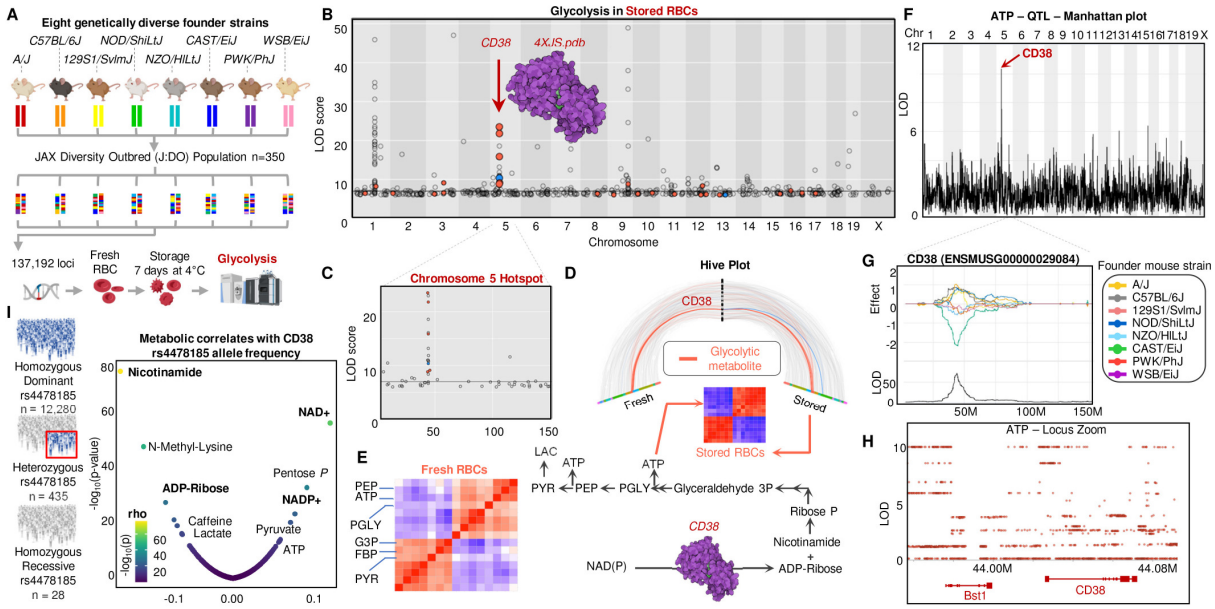
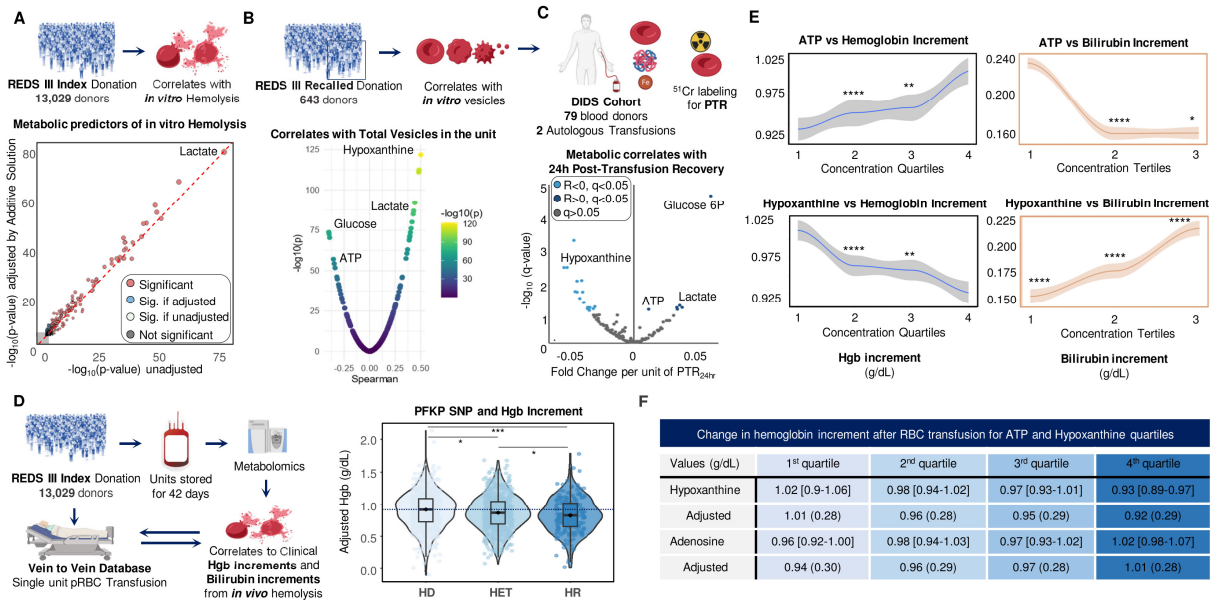


Figure 3 – Association between glycolysis and CD38 polymorphisms in fresh and stored murine RBCs from the Diversity Outbred population. Metabolomics analysis was performed in fresh and stored RBCs from 350 mice derived from extensive cross-breeding (>40 generations) of 8 genetically distinct founder strains (Jackson Laboratory's Diversity Outbred – J:DO – **A**). A hot spot of mQTL were identified on chromosome 5 (**B**), which included many glycolytic metabolites (**C**), as further emphasized with by the hive plot in (**D**) and heat map of metabolic correlates to this locus in fresh and stored RBCs (**E**). Manhattan plot for ATP mQTL analysis (**F**) confirms the association with a polymorphic region on chromosome 5 (**G**), which codes for CD38 (**H**) and associates with all glycolytic and pentose phosphate pathway (PPP) metabolites in stored murine RBCs (heat map in indent in **D**, extensive in **Supporting figure 8**). The SNP rs4478185 was identified as the top CD38 SNPs associated with glycolysis and PPP metabolites in human RBCs from the REDS RBC Omics Index cohort confirming observations in mice (**I**).



Change in hemoglobin increment after RBC transfusion for ATP and Hypoxanthine quartiles				
Values (g/dL)	1st quartile	2nd quartile	3rd quartile	4th quartile
Hypoxanthine	1.02 [0.9-1.06]	0.98 [0.94-1.02]	0.97 [0.93-1.01]	0.93 [0.89-0.97]
Adjusted	1.01 (0.28)	0.96 (0.28)	0.95 (0.29)	0.92 (0.29)
Adenosine	0.96 [0.92-1.00]	0.98 [0.94-1.03]	0.97 [0.93-1.02]	1.02 [0.98-1.07]
Adjusted	0.94 (0.30)	0.96 (0.29)	0.97 (0.28)	1.01 (0.28)

Figure 4 – Positive and negative association of glycolysis/ATP and hypoxanthine to in vitro and in vivo hemolysis in healthy and critically ill individuals receiving autologous and non-autologous transfusion, respectively. End of storage spontaneous hemolysis - one of the two FDA gold standards for the quality of stored pRBC products - was measured in all 13,029 units. Lactate levels at storage day 42 were the top predictor of spontaneous hemolysis (**A**). In **B,** metabolic correlates to the degree of vesiculation in 643 units (1,929 samples at storage day 10, 23 and 42) from the REDS RBC Omics Recalled donor population (i.e., a subset of donors with extreme hemolytic propensity – 5th and 95th percentile of the total Index population). In a separate cohort of 79 healthy volunteer donations and autologous transfusion events, glycolytic metabolites and ATP levels were positively associated to, and hypoxanthine negatively associated to measurements of post-transfusion recovery with ^{51}Cr in healthy volunteers – the other FDA gold standard for monitoring the quality of stored RBCs (**C**). By accessing the "vein-to-vein" database of the REDS RBC Omics program, we could link blood donor PFKP rs2388595 status (homozygous dominant = HD n=521; heterozygous = HET n=576; homozygous recessive = HR n=204) to adjusted hemoglobin increments in clinical recipients of irradiated single pRBC unit transfusions (**D**). Similarly, we linked end of storage ATP and hypoxanthine levels to hemoglobin increments (n=5,816 records available in the database) and bilirubin increments (n=914) upon heterologous transfusion of products from the same donors to critically ill recipients requiring transfusion (**E-F**).