

Comment on *Kammers et al*, page 959

eQTLs in platelets and iPSC-megakaryocytes

Elizabeth A. Middleton and Jesse W. Rowley | The University of Utah

In this issue of *Blood*, the study by Kammers et al¹ identifies genetic variants associated with gene expression in 290 platelet and 185 induced pluripotent stem cell (iPSC) megakaryocyte (MK) samples to fill in missing links between genotype and platelet phenotypes.

Platelets are unique, anucleate, effector cells with roles in immunity, wound healing, hemostasis, and thrombosis. Heritable platelet traits, such as mean platelet volume, platelet count, and their propensity to aggregate and form clots, have been linked to disease.^{2,3} Platelet traits have also been associated with differences in platelet gene expression.^{4,5} Platelet gene expression varies between individuals, is heritable, and is highly repeatable for example, stably retained over time.⁶ Genetic variants associated with gene expression, called expression quantitative trait loci (eQTL) (see figure), account for a significant portion of this stable variation between individuals.^{6,7} A previous study by Simon et al of platelet gene expression in 154 European American and African American subjects identified eQTL for 612 genes (called eGenes), many of which were unique to platelets.⁷

However, eQTLs found for platelets alone may not account for all of the inherited variation in platelets. Anucleate platelets inherit much of their RNA and protein content from their parent MKs. From transcription in the MK, followed by loading into platelets and 7 to 10 days of circulation, differences in RNA levels inevitably accumulate (differential RNA sorting into platelets, turnover, uptake), whereas their protein products and functional consequences might remain. Therefore, MKs could harbor novel gene-variant associations that are lost at the RNA level in platelets yet are still responsible for differences in platelet proteins and function. Unfortunately, native MKs capable of forming platelets that reside in the bone marrow and lung⁸ are rare and difficult to access without invasive procedures to acquire sufficient numbers for eQTL analysis. To address this gap, as outlined in the figure, Kammer et al generated

iPSC-derived MKs from 185 African American and European American donors and compared their transcriptome to platelets from 290 donors. Utilizing RNA-seq on polyadenylated RNA (coding transcripts), the authors found that 91.3% of the genes expressed in platelets were expressed in iPSC-MKs, whereas only 53.1% of the iPSC-MK-expressed genes were in platelets, with a modest Spearman correlation between shared genes of 0.46.

Comparing RNA levels with DNA variants called by genome sequencing, the authors identified 1830 eQTLs in platelets. This replicated 85% of previously identified platelet eQTLs⁷ and adds nearly 3 times as many new eQTLs. Also identified were 946 eQTLs for MKs. Of these, 323 matching genes were identified with significant eQTLs in both MKs and platelets, whereas 57 of these shared the same lead eQTL. This number is surprisingly small since the bulk of RNA in platelets presumably comes from their parent MKs. However, considering donor differences and that the directionality of effect was predominantly the same for those eQTLs that reached statistical significance in 1 tissue but not the other, the actual regulatory overlap is potentially larger.

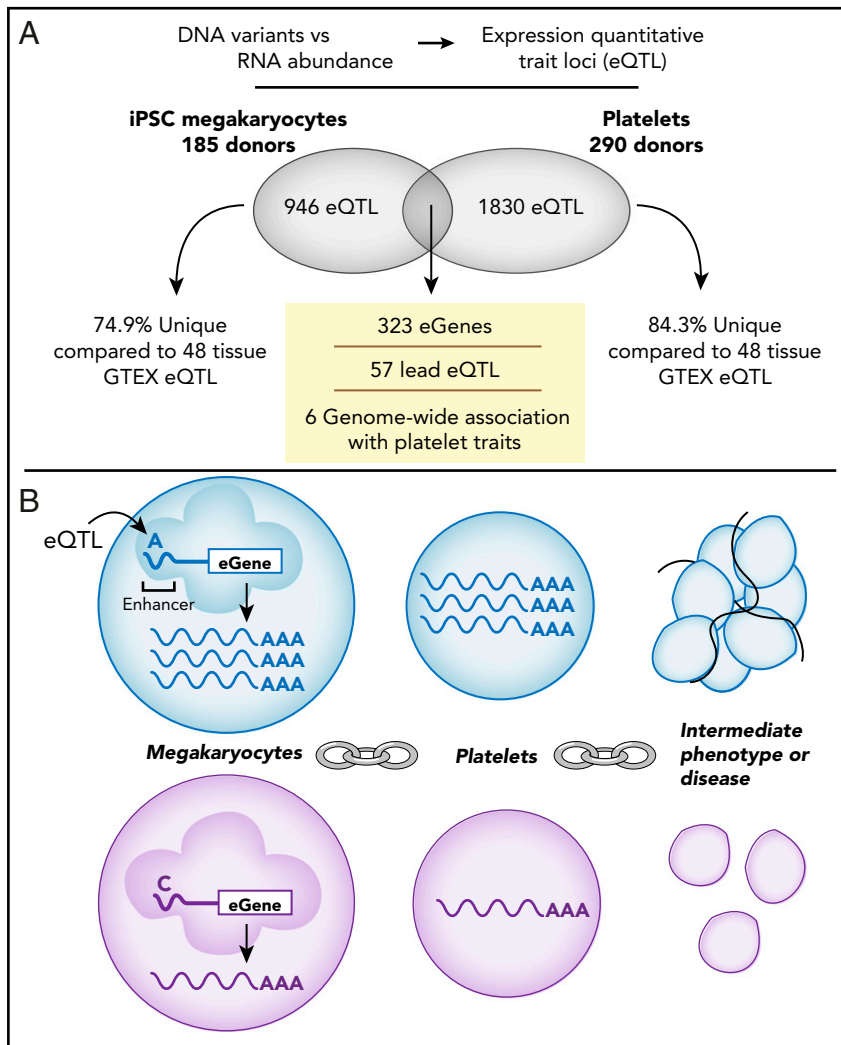
Variants coded within the RNA itself may affect transcription, structure, stability, localization, or the individual amino acid sequence of the protein. However, how do eQTLs found in noncoding regions distant to the gene body influence RNA expression? The authors found, for both platelets and MKs, an enrichment for lead eQTLs in regulatory regions previously identified in MKs that include long-range interactions such as DNase-sensitive enhancer elements. This points to eQTL-

mediated transcriptional regulation initiated in MKs and carried over into platelets.

Genome-wide association studies (GWAS) have found DNA variants associated with thousands of diseases, including those where platelets participate, such as cardiovascular disease.^{9,10} However, the relevance of these variants has been difficult to interpret. For one, DNA variants at neighboring genetic loci are co-inherited (linkage disequilibrium), so the exact responsible variant among neighbors is difficult to define. In addition, between DNA at one end and disease outcome at the other is a chain of complex molecular, cellular, intercellular, and organismal interactions. Therefore, recent studies have emphasized linking DNA variants with more narrowly defined molecular and cellular intermediates, like gene expression, that in turn can be ultimately linked to disease. Databases of eQTLs found for many different tissues have helped link GWAS variants to gene expression, but the regulatory effect of eQTLs is often tissue specific; thus eQTLs are not found in current tissue databases for many variants linked with platelet traits. Consistent with this, the authors found a low overlap with platelet or MK lead eQTLs (the eQTL with the strongest association with a gene) and those previously found for 48 other tissues, suggesting platelet- and MK-specific regulation of gene expression.

To identify possible new links between DNA variants, gene expression, and platelet phenotypes, the authors compared their MK and platelet eQTLs with previously published GWAS studies of platelet traits. Of 1065 significant GWAS variants for platelet traits, 24 matched eQTLs in either platelets or MKs, and 6 matched in both. Additional eQTLs in linkage disequilibrium were also found for other GWAS single-nucleotide polymorphisms. The identification of these eQTLs provides a provocative intermediate link between DNA variants and GWAS phenotype (see figure panel B), an additional step toward understanding the functional consequences of a given genotype.

There are some caveats that should be considered when interpreting the study. The iPSC-derived populations of cells contained ~60% of cells double positive for the MK markers CD41 and CD42a. Although there was significant overlap with the iPSC-MKs and platelet gene expression, substantial differences in gene



(A) Schematic representation of eQTLs identified in iPSC-derived MKs ($n = 185$) and platelets ($n = 290$) from healthy donors. Of eQTLs identified in MKs and platelets, 74.9% and 84.3%, respectively, were unique to tissue type. Three hundred twenty-three eGenes and 57 matching lead *cis*-eQTLs were in both MKs and platelets, and 6 map to GWAS identified platelet traits. (B) eQTLs are genetic variants, or differences in DNA between individuals (depicted by the A [blue] vs C [purple] nucleotides) that are associated with RNA abundance. eQTLs in both platelets and MKs were enriched in MK regulatory regions, such as enhancers linking differential regulation of transcription in MKs to RNA abundance carried over into platelets. These differences may in turn impact proteins, phenotypes, and function in health and disease.

expression will unquestionably remain between in vivo MKs and an impure population of in vitro-derived MKs. Because many eQTLs are tissue and context dependent, some of the eQTLs found for iPSC-MKs may be specific for in vitro cells.

Nevertheless, the identified MK variants may prove useful when choosing which GWAS variants to rule in or out for follow-up studies to establish causation. Establishing causation with functional experimentation is an essential next step to deciphering

gene-variant functions that will ultimately inform mechanisms of platelet function and forge new pathways toward disease prevention and treatment.

Conflict-of-interest disclosure: The authors declare no competing financial interests. ■

REFERENCES

1. Kammers K, Taub MA, Rodriguez B, et al. Transcriptional profile of platelets and iPSC-derived megakaryocytes from whole-genome and RNA sequencing. *Blood*. 2021;137(7):959-968.
2. Eicher JD, Lettre G, Johnson AD. The genetics of platelet count and volume in humans. *Platelets*. 2018;29(2):125-130.
3. Puurunen MK, Hwang S-J, Larson MG, et al. ADP platelet hyperreactivity predicts cardiovascular disease in the FHS (Framingham Heart Study). *J Am Heart Assoc*. 2018;7(5):e008522.
4. Goodall AH, Burns P, Salles I, et al; Bloodomics Consortium. Transcription profiling in human platelets reveals LRRFIP1 as a novel protein regulating platelet function. *Blood*. 2010;116(22):4646-4656.
5. Edelstein LC, Simon LM, Montoya RT, et al. Racial differences in human platelet PAR4 reactivity reflect expression of PCTP and miR-376c. *Nat Med*. 2013;19(12):1609-1616.
6. Rondina MT, Voora D, Simon LM, et al. Longitudinal RNA-seq analysis of the repeatability of gene expression and splicing in human platelets identifies a platelet selp splice QTL. *Circ Res*. 2020;126(4):501-516.
7. Simon LM, Chen ES, Edelstein LC, et al. Integrative multi-omic analysis of human platelet eQTLs reveals alternative start site in mitofusin 2. *Am J Hum Genet*. 2016;98(5):883-897.
8. Lefrançois E, Ortiz-Muñoz G, Caudrillier A, et al. The lung is a site of platelet biogenesis and a reservoir for haematopoietic progenitors. *Nature*. 2017;544(7648):105-109.
9. Erdmann J, Kessler T, Munoz Venegas L, Schunkert H. A decade of genome-wide association studies for coronary artery disease: the challenges ahead. *Cardiovasc Res*. 2018;114(9):1241-1257.
10. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 2019;20(8):467-484.

DOI 10.1182/blood.202009461

© 2021 by The American Society of Hematology