Genome-wide association studies of hematologic phenotypes: a window into human hematopoiesis

Vijay G. Sankaran^{1,2,3} and Stuart H. Orkin¹

¹Division of Hematology/Oncology, Boston Children's Hospital and Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115

²Broad Institute, Cambridge, MA 02142

³Whitehead Institute for Biomedical Research, Cambridge, MA 02142

Abstract

The study of human hematopoiesis is often limited by the inability to manipulate this process *in vivo* and differences that exist between humans and commonly employed model organisms. However, human genetics provides a way to gain insight into natural variation in a variety of hematologic phenotypes and creates an opportunity to better understand hematopoiesis. In this review, we discuss how genome-wide association studies are revealing common genetic variation that is associated with hematologic traits and diseases. We discuss how the resulting insight from these studies promises to increase our understanding of human hematopoiesis and outline the challenges that lay ahead in this field.

Introduction

Hematopoiesis is among the best-understood paradigms for cell differentiation [1]. Studies in this field have traditionally been driven by the use of powerful reverse genetic approaches in a variety of model systems. However, there are important differences between clinically relevant aspects of hematopoiesis in humans and that of commonly used model organisms, including mice. For example, hemoglobin gene regulation is divergent between humans and mice [2,3]. Additionally, mutations in humans that result in specific hematologic phenotypes may not be faithfully recapitulated in mice [4-7]. Unless appropriate studies in humans are undertaken, such important differences may not be uncovered. However, the analysis of human hematopoiesis is hampered by the inability to perform *in vivo* manipulations, as can be readily accomplished in model organisms. Here we discuss how studies of common genetic variation in hematologic traits and diseases provide a window with which to explore aspects of human hematopoiesis *in vivo*. When coupled with studies in model systems, these approaches offer an opportunity to reveal new regulatory processes critical for the maintenance of the blood system.

The hematopoietic system is composed of three major cell types: red blood cells (erythrocytes) that provide oxygen carrying capacity to the circulation, platelets that are

critical for hemostasis through blood clot formation, and white blood cells (primarily neutrophils, monocytes, and lymphocytes) that maintain host defense to infections [1]. Each day, a healthy human produces over 10^{11} new red blood cells (erythrocytes) and platelets, as well as 10^{10} neutrophils. If this output is not maintained, the individual is at risk for a number of hematologic disorders. Because of this rapid turnover in circulating blood cells, the hematopoietic system is poised to continually replenish these lineages through the proliferation, differentiation, and maturation of immature progenitor populations that are the progeny of self-renewing stem cells (Figure 1) [1].

Over the past decade, the search for the genetic basis of complex traits and disease has been enabled through the use of genome-wide association studies (GWAS) [8]. In such studies, hundreds of thousands to millions of polymorphic markers across the genome are assessed to enable searches for statistical association with either a disease-state or with a quantitative trait [9,10]. Information about the linkage disequilibrium pattern for various polymorphisms has allowed comprehensive searches for common variants in the genome associated with a disease or trait of interest. Here we review such studies that have been performed for various hematologic parameters and diseases and discuss how these findings are enabling new insight into human hematopoiesis.

The genetics of erythroid phenotypes

The earliest GWAS studies of hematologic phenotypes were focused on fetal hemoglobin (HbF) levels [11,12]. The interest in this trait stems from the ability of increased HbF to ameliorate the clinical symptoms of the two major disorders of hemoglobin, sickle cell disease (SCD) and β -thalassemia [13]. These GWAS studies revealed three major loci associated with HbF levels, which collectively explained nearly half of the variation in this trait – a contrast to the majority of other traits and diseases examined by GWAS [8,14]. The loci identified included the β -globin gene cluster on chromosome 11, as well as variants on chromosome 6 in a region between the genes *HBSIL* and *MYB* and variants within the gene *BCL11A* on chromosome 2 [14,15]. The elucidation of HbF-associated genetic variants in the *BCL11A* gene led to work showing that the product of this gene is a major transcriptional regulator of the physiologic fetal-to-adult hemoglobin switch and HbF silencing in humans [16]. More recent studies have shown that ablation of *Bcl11a* leads to substantial reactivation of HbF expression and amelioration of SCD phenotypes in mouse models [17], supporting the findings from studies of HbF-associated genetic variants in SCD patients [15]. Other studies have suggested that the *MYB* gene may have an important role in the regulation of HbF levels [18,19], but the exact mechanisms underlying these observations and how precisely this relates to the HbF-associated variants on chromosome 6 remains unclear [20]. It should be noted that the variants in the *HBSIL-MYB* intergenic region are pleiotropically associated with several other hematologic traits (including WBC and platelet counts), albeit to a far lesser extent than their effect on HbF levels [21]. These studies have revealed promising therapeutic targets for HbF induction in patients with β -thalassemia and SCD. Indeed, GWAS of clinical severity in β -thalassemia patient populations found that the HbF-associated loci are significantly associated with and may explain the majority of variation in the clinical course [22,23]. It should be noted that all the GWAS of HbF levels have been performed on adults and while the three loci that were uncovered may have

important roles in silencing of HbF, other common genetic variants may influence the developmental fetal-to-adult globin switch.

Since the initial studies on HbF levels, further studies have examined a variety of other clinically-relevant erythroid parameters including hemoglobin levels, hematocrit (the percentage of the blood volume occupied by erythrocytes), red blood cell count, and red cell size (mean corpuscular volume, MCV) [24-32]. In these other erythrocyte phenotype studies, the loci reaching genome-wide significance collectively explain < 10% and often only a few percent of the total variation in those traits. These studies have implicated > 20 loci in erythrocyte traits. However, with a few exceptions, the underlying molecular mechanisms remain unclear. The *BCL11A* and *HBS1L-MYB* loci have been implicated in the regulation of erythrocyte size and this may be related to the same molecular mechanisms that lead to variation in HbF levels [24]. The *TMPRSS6* locus has been strongly associated with a variety of erythroid traits, most notably hemoglobin levels, in a number of different GWAS [24-27]. Interestingly, rare mutations in the *TMPRSS6* gene, which encodes a transmembrane serine protease expressed in the liver that is necessary for the regulation of iron absorption, are associated with iron deficiency states [33]. This is an excellent example of how GWAS can highlight informative biology. In addition, some of these variants may be important modifiers of common forms of anemia, though this remains to be formally tested [34].

The mechanisms underlying the loci implicated in the erythrocyte trait GWAS have been largely unexplored. However, using clues from these studies and by performing mechanistic follow up using both human and mouse erythroid cells, it was found that the *CCND3* gene on chromosome 6 (encoding cyclin D3) plays a critical role as a regulator of the number of divisions that occur during terminal erythropoiesis, thereby regulating erythrocyte size and number [35]. In addition to providing important insight into how normal erythropoiesis proceeds, these findings may have implications for understanding pathologic states where erythropoiesis is disrupted that are often characterized by changes in erythrocyte size. Moreover, these findings could also have important implications for ongoing work to recapitulate the process of erythropoiesis *ex vivo* for potential use in transfusions [36].

Common genetic variation in platelet phenotypes

Platelets are critical for the formation of blood clots and are continually produced from multinuclear precursors in the bone marrow known as megakaryocytes (Figure 1). In clinical practice, both platelet count and size (mean platelet volume, MPV) are measured routinely. Given the association of platelets with the formation of blood clots, there has been great interest in examining these traits, which can predispose individuals to cardiovascular events, such as myocardial infarctions [21,37,38]. A number of GWAS have examined both platelet count and MPV in humans [25,31,32,37,39-41]. A large meta-analysis of these GWAS, involving > 66,000 individuals, has recently been reported and has identified 68 loci associated with platelet phenotypes, including 53 newly discovered loci [37]. This study highlights the value of performing large-scale meta-analyses to uncover new loci associated with particular phenotypes. While the effect sizes of the majority of uncovered variants are small (generally < 1% of the variation), many of the genes in these loci appear critical for

aspects of megakaryopoiesis and platelet formation [37]. However, while genomic approaches have suggested that many of these genes are highly expressed in megakaryocytes and studies in zebrafish and *Drosophila* models of hematopoiesis support an important role for these genes, the precise molecular mechanisms underlying the association between platelet phenotypes and the variants in these loci remain unclear [37]. Some studies have begun to explore the relationship between specific platelet-associated variants and their effect on the expression of specific genes, such as *PIK3CG* [42] – however, even in this well-studied case, the mechanism by which the *PIK3CG* gene influences platelet formation remains unclear.

The landscape of common genetic variation in WBC counts and diseases

WBCs are critical for the defense against foreign microorganisms and compose both the innate and acquired immune systems [1]. In addition to their role in normal immunity, the development of these cells often goes awry in human conditions including autoimmunity and hematologic malignancies. As with other hematologic phenotypes, WBC phenotypes have also been the subject of GWAS [25,28,31,32,43-49]. In contrast to erythrocytes and platelets, WBCs constitute a heterogeneous population of cells. The majority of circulating WBCs are neutrophils (50-70% of WBCs in healthy individuals), which form an early defense against invading microorganisms. It is therefore not surprising that there is a significant overlap between loci that were implicated in variation of both the total WBC and neutrophil counts [21,46-48]. For example the *PSMD3-CSF3* locus on chromosome 17 has been associated with both WBC and neutrophil counts in several studies. This association is interesting, since a strong candidate for these traits is the *CSF3* gene, which encodes the granulocyte colony stimulating factor (G-CSF) that is used therapeutically to increase neutrophil output. Further work is needed to examine whether these variants may influence G-CSF levels in humans and whether this may also affect the response to G-CSF therapy [50]. Studies have also suggested the role of loci associated with monocyte levels [28,46-49], including the *ITGA4* gene locus on chromosome 2 encoding the $\alpha 4$ integrin subunit that is necessary for monocyte migration [51]. Given their role in common allergic diseases and asthma, there has also been interest in understanding genetic variants that affect the levels of eosinophils and basophils. Some interesting candidate associations have been uncovered, including variants at the *GATA2* and *ERG* loci [48]. Interestingly, several of the loci associated with eosinophil count were also associated with allergic asthma [43], supporting the recently uncovered relationship between eosinophils and asthma [52].

While some promising associations for neutrophil, monocyte, eosinophil, and basophil counts have been reported, the genetic studies of lymphocyte counts have yielded far less success [46-48]. This may be attributable to the fact that most lymphocyte counts do not differentiate between different subsets that require strikingly different molecular programs for their development [1]. Indeed, a study of CD4:CD8 ratios of T lymphocytes has yielded some promising association signals [44], suggesting that such approaches on a larger scale may be beneficial. In several instances however, studies of autoimmune diseases have revealed important regulators of particular lymphocyte subsets [53,54]. In addition, studies of B lymphocyte malignancies including chronic lymphocytic leukemia (CLL) and acute lymphoblastic leukemia (ALL) have revealed loci that contain genes known to be critical for

normal B lymphocyte development [55-59]. These findings collectively suggest that common variation in the normal development of these lymphocyte populations may confer susceptibility to human disease and hence molecular studies of such processes in humans may be vital to develop better therapeutic approaches for these disorders.

Challenges and future directions

While the past few years has witnessed an impressive output of data on common genetic variation in human hematopoiesis, four major challenges arise from this work. **(1)** As illustrated by the example on platelet phenotypes, larger meta-analyses of GWAS can uncover new aspects of biology that are not easily appreciated [37]. Although many of the loci uncovered from such studies can have rather small effect sizes, the importance of the genes involved in these loci may be critical for understanding important aspects of hematopoiesis [35]. It is important to remember that the genetic contribution to variation seen in GWAS at any locus provides only a minimal estimate of the biological contribution and major perturbations of these genes can show dramatic phenotypes. **(2)** While the loci that have been uncovered from GWAS studies provide a “parts list” of potential regions associated with a particular trait, the genes underlying these effects remain to be uncovered. Occasionally, nearest gene approaches are used to implicate a specific gene in a trait, but more unbiased and rigorous genomic approaches should be applied to properly nominate candidate genes in these loci and uncover causal variants [10,60]. If such mapping approaches are coupled to other orthogonal genomic datasets, such as cell-type specific expression or functional datasets, insight into likely candidate genes can be gained [10,37]. **(3)** While broader analyses will be useful in annotating the genes that are involved in the various traits discussed above, detailed molecular and mechanistic studies are needed to better understand how these genes play a role in various aspects of hematopoiesis. Even for well-studied cases [16,35], further work is needed to better understand the mechanisms involved. **(4)** Finally, it will be important to integrate the information from common variant studies with rare mutation analyses and studies of genes involved in hematologic disorders. Rare variant studies of hematologic traits are just now being reported and will likely have improved power to detect new variants in the ensuing years [61]. As illustrated by the studies of hematologic malignancies and autoimmune conditions, important lessons can be learned about susceptibility to these diseases by studying aspects of normal human hematopoiesis [53,56]. As new studies surmount these challenges in the coming years, our understanding of human hematopoiesis will undoubtedly improve significantly.

Acknowledgements

We thank L. Ludwig, R. Do, C. Walkley, J. Flygare, and J. Eng for helpful discussions. We are grateful to T. DiCesare for assistance in preparing figures. V.G.S. is supported by NIH grant T32 HL007574-30. S.H.O. is an Investigator of the HHMI and is supported by NIH grants HL32259 and HL32262.

References and recommended reading

1. Orkin SH, Zon LI. Hematopoiesis: An evolving paradigm for stem cell biology. *Cell*. 2008; 132(4): 631–644. [PubMed: 18295580]

2. Sankaran VG, Xu J, Ragoczy T, Ippolito GC, Walkley CR, Maika SD, Fujiwara Y, Ito M, Groudine M, Bender MA, Tucker PW, et al. Developmental and species-divergent globin switching are driven by *bcl11a*. *Nature*. 2009; 460(7259):1093–1097. [PubMed: 19657335]
3. McGrath KE, Frame JM, Fromm GJ, Koniski AD, Kingsley PD, Little J, Bulger M, Palis J. A transient definitive erythroid lineage with unique regulation of the beta-globin locus in the mammalian embryo. *Blood*. 2011; 117(17):4600–4608. [PubMed: 21378272]
4. Sankaran VG, Ghazvinian R, Do R, Thiru P, Vergilio JA, Beggs AH, Sieff CA, Orkin SH, Nathan DG, Lander ES, Gazda HT. Exome sequencing identifies *gata1* mutations resulting in diamond-blackfan anemia. *J Clin Invest*. 2012; 122(7):2439–2443. [PubMed: 22706301]
5. Hollanda LM, Lima CS, Cunha AF, Albuquerque DM, Vassallo J, Ozelo MC, Joazeiro PP, Saad ST, Costa FF. An inherited mutation leading to production of only the short isoform of *gata-1* is associated with impaired erythropoiesis. *Nat Genet*. 2006; 38(7):807–812. [PubMed: 16783379]
6. Sankaran VG, Xu J, Byron R, Greisman HA, Fisher C, Weatherall DJ, Sabath DE, Groudine M, Orkin SH, Premawardhena A, Bender MA. A functional element necessary for fetal hemoglobin silencing. *N Engl J Med*. 2011; 365(9):807–814. [PubMed: 21879898]
7. Tao J, Zhu M, Wang H, Afelik S, Vasievich MP, Chen XW, Zhu G, Jensen J, Ginsburg D, Zhang B. *Sec23b* is required for the maintenance of murine professional secretory tissues. *Proc Natl Acad Sci U S A*. 2012; 109(29):E2001–2009. [PubMed: 22745161]
8. Lander ES. Initial impact of the sequencing of the human genome. *Nature*. 2011; 470(7333):187–197. [PubMed: 21307931]
9. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science*. 2008; 322(5903):881–888. [PubMed: 18988837]
10. Raychaudhuri S. Mapping rare and common causal alleles for complex human diseases. *Cell*. 2011; 147(1):57–69. [PubMed: 21962507]
11. Uda M, Galanello R, Sanna S, Lettre G, Sankaran VG, Chen W, Usala G, Busonero F, Maschio A, Albai G, Piras MG, et al. Genome-wide association study shows *bcl11a* associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. *Proc Natl Acad Sci U S A*. 2008; 105(5):1620–1625. [PubMed: 18245381]
12. Menzel S, Garner C, Gut I, Matsuda F, Yamaguchi M, Heath S, Foglio M, Zelenika D, Boland A, Rooks H, Best S, et al. A qtl influencing f cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat Genet*. 2007; 39(10):1197–1199. [PubMed: 17767159]
13. Sankaran VG, Nathan DG. Reversing the hemoglobin switch. *N Engl J Med*. 2010; 363(23):2258–2260. [PubMed: 21121839]
14. Galarneau G, Palmer CD, Sankaran VG, Orkin SH, Hirschhorn JN, Lettre G. Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. *Nat Genet*. 2010; 42(12):1049–1051. [PubMed: 21057501]
15. Lettre G, Sankaran VG, Bezerra MA, Araujo AS, Uda M, Sanna S, Cao A, Schlessinger D, Costa FF, Hirschhorn JN, Orkin SH. DNA polymorphisms at the *bcl11a*, *hbs11-myb*, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci U S A*. 2008; 105(33):11869–11874. [PubMed: 18667698]
16. Sankaran VG, Menne TF, Xu J, Akie TE, Lettre G, Van Handel B, Mikkola HK, Hirschhorn JN, Cantor AB, Orkin SH. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor *bcl11a*. *Science*. 2008; 322(5909):1839–1842. [PubMed: 19056937]
17. Xu J, Peng C, Sankaran VG, Shao Z, Esrick EB, Chong BG, Ippolito GC, Fujiwara Y, Ebert BL, Tucker PW, Orkin SH. Correction of sickle cell disease in adult mice by interference with fetal hemoglobin silencing. *Science*. 2011; 334(6058):993–996. [PubMed: 21998251]
18. Sankaran VG, Menne TF, Scepanovic D, Vergilio JA, Ji P, Kim J, Thiru P, Orkin SH, Lander ES, Lodish HF. *Microrna-15a* and *-16-1* act via *myb* to elevate fetal hemoglobin expression in human trisomy 13. *Proc Natl Acad Sci U S A*. 2011; 108(4):1519–1524. [PubMed: 21205891]
19. Jiang J, Best S, Menzel S, Silver N, Lai MI, Surdulescu GL, Spector TD, Thein SL. *Cmyb* is involved in the regulation of fetal hemoglobin production in adults. *Blood*. 2006; 108(3):1077–1083. [PubMed: 16861354]
20. Stadhouders R, Thongjuea S, Andrieu-Soler C, Palstra RJ, Bryne JC, van den Heuvel A, Stevens M, de Boer E, Kockx C, van der Sloot A, van den Hout M, et al. Dynamic long-range chromatin

- interactions control myb proto-oncogene transcription during erythroid development. *Embo J*. 2012; 31(4):986–999. [PubMed: 22157820]
21. Okada Y, Kamatani Y. Common genetic factors for hematological traits in humans. *J Hum Genet*. 2012; 57(3):161–169. [PubMed: 22277899]
 22. Nuinon M, Makarasara W, Mushiroda T, Setianingsih I, Wahidiyat PA, Sripichai O, Kumasaka N, Takahashi A, Svasti S, Munkongdee T, Mahasirimongkol S, et al. A genome-wide association identified the common genetic variants influence disease severity in beta0-thalassemia/hemoglobin e. *Hum Genet*. 2010; 127(3):303–314. [PubMed: 20183929]
 23. Galanello R, Sanna S, Perseu L, Sollaino MC, Satta S, Lai ME, Barella S, Uda M, Usala G, Abecasis GR, Cao A. Amelioration of sardinian beta0 thalassemia by genetic modifiers. *Blood*. 2009; 114(18):3935–3937. [PubMed: 19696200]
 24. Ganesh SK, Zakai NA, van Rooij FJ, Soranzo N, Smith AV, Nalls MA, Chen MH, Kottgen A, Glazer NL, Dehghan A, Kuhnle B, et al. Multiple loci influence erythrocyte phenotypes in the charge consortium. *Nat Genet*. 2009; 41(11):1191–1198. [PubMed: 19862010]
 25. Soranzo N, Spector TD, Mangino M, Kuhnle B, Rendon A, Teumer A, Willenborg C, Wright B, Chen L, Li M, Salo P, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the haemgen consortium. *Nat Genet*. 2009; 41(11):1182–1190. [PubMed: 19820697]
 26. Chambers JC, Zhang W, Li Y, Sehmi J, Wass MN, Zabaneh D, Hoggart C, Bayele H, McCarthy MI, Peltonen L, Freimer NB, et al. Genome-wide association study identifies variants in *tmprss6* associated with hemoglobin levels. *Nat Genet*. 2009; 41(11):1170–1172. [PubMed: 19820698]
 27. Benyamin B, Ferreira MA, Willemsen G, Gordon S, Middelberg RP, McEvoy BP, Hottenga JJ, Henders AK, Campbell MJ, Wallace L, Frazer IH, et al. Common variants in *tmprss6* are associated with iron status and erythrocyte volume. *Nat Genet*. 2009; 41(11):1173–1175. [PubMed: 19820699]
 28. Ferreira MA, Hottenga JJ, Warrington NM, Medland SE, Willemsen G, Lawrence RW, Gordon S, de Geus EJ, Henders AK, Smit JH, Campbell MJ, et al. Sequence variants in three loci influence monocyte counts and erythrocyte volume. *Am J Hum Genet*. 2009; 85(5):745–749. [PubMed: 19853236]
 29. Ding K, Shameer K, Jouni H, Masys DR, Jarvik GP, Kho AN, Ritchie MD, McCarty CA, Chute CG, Manolio TA, Kullo IJ. Genetic loci implicated in erythroid differentiation and cell cycle regulation are associated with red blood cell traits. *Mayo Clin Proc*. 2012; 87(5):461–474. [PubMed: 22560525]
 30. Kullo IJ, Ding K, Jouni H, Smith CY, Chute CG. A genome-wide association study of red blood cell traits using the electronic medical record. *PLoS One*. 2010; 5(9)
 31. Lo KS, Wilson JG, Lange LA, Folsom AR, Galarneau G, Ganesh SK, Grant SF, Keating BJ, McCarroll SA, Mohler ER 3rd, O'Donnell CJ, et al. Genetic association analysis highlights new loci that modulate hematological trait variation in caucasians and african americans. *Hum Genet*. 2011; 129(3):307–317. [PubMed: 21153663]
 32. Kamatani Y, Matsuda K, Okada Y, Kubo M, Hosono N, Daigo Y, Nakamura Y, Kamatani N. Genome-wide association study of hematological and biochemical traits in a japanese population. *Nat Genet*. 2010; 42(3):210–215. [PubMed: 20139978]
 33. Finberg KE, Heeney MM, Campagna DR, Aydinok Y, Pearson HA, Hartman KR, Mayo MM, Samuel SM, Strouse JJ, Markianos K, Andrews NC, et al. Mutations in *tmprss6* cause iron-refractory iron deficiency anemia (irida). *Nat Genet*. 2008; 40(5):569–571. [PubMed: 18408718]
 34. Balarajan Y, Ramakrishnan U, Ozaltin E, Shankar AH, Subramanian SV. Anaemia in low-income and middle-income countries. *Lancet*. 2011; 378(9809):2123–2135. [PubMed: 21813172]
 - 35••. Sankaran VG, Ludwig LS, Sicinska E, Xu J, Bauer DE, Eng JC, Patterson HC, Metcalf RA, Natkunam Y, Orkin SH, Sicinski P, et al. Cyclin d3 coordinates the cell cycle during differentiation to regulate erythrocyte size and number. *Genes Dev*. 2012; 26(18):2075–2087. [PubMed: 22929040] [This article provides an example of how follow-up from GWAS can help identify new molecular mechanisms that are critical for hematopoiesis. In this article, cyclin D3 is identified as an important regulator of the number of divisions that occur during terminal erythropoiesis, thereby controlling erythrocyte size and number.]

36. Migliaccio AR, Whitsett C, Papayannopoulou T, Sadelain M. The potential of stem cells as an in vitro source of red blood cells for transfusion. *Cell Stem Cell*. 2012; 10(2):115–119. [PubMed: 22305561]
- 37•. Gieger C, Radhakrishnan A, Cvejic A, Tang W, Porcu E, Pistis G, Serbanovic-Canic J, Elling U, Goodall AH, Labrune Y, Lopez LM, et al. New gene functions in megakaryopoiesis and platelet formation. *Nature*. 2011; 480(7376):201–208. [PubMed: 22139419] [This impressive meta-analysis of > 66,000 individuals illustrates the value of large-scale GWAS analyses to identify new regulators of hematopoiesis and also illustrates the use of orthogonal genomic datasets to narrow down candidate genes within the loci identified.]
38. Martin JF, Kristensen SD, Mathur A, Grove EL, Choudry FA. The causal role of megakaryocyte-platelet hyperactivity in acute coronary syndromes. *Nat Rev Cardiol*. 2012
39. Meisinger C, Prokisch H, Gieger C, Soranzo N, Mehta D, Rosskopf D, Lichtner P, Klopp N, Stephens J, Watkins NA, Deloukas P, et al. A genome-wide association study identifies three loci associated with mean platelet volume. *Am J Hum Genet*. 2009; 84(1):66–71. [PubMed: 19110211]
40. Soranzo N, Rendon A, Gieger C, Jones CI, Watkins NA, Menzel S, Doring A, Stephens J, Prokisch H, Erber W, Potter SC, et al. A novel variant on chromosome 7q22.3 associated with mean platelet volume, counts, and function. *Blood*. 2009; 113(16):3831–3837. [PubMed: 19221038]
41. Qayyum R, Snively BM, Ziv E, Nalls MA, Liu Y, Tang W, Yanek LR, Lange L, Evans MK, Ganesh S, Austin MA, et al. A meta-analysis and genome-wide association study of platelet count and mean platelet volume in african americans. *PLoS Genet*. 2012; 8(3):e1002491. [PubMed: 22423221]
42. Paul DS, Nisbet JP, Yang TP, Meacham S, Rendon A, Hautaviita K, Tallila J, White J, Tijssen MR, Sivapalaratnam S, Basart H, et al. Maps of open chromatin guide the functional follow-up of genome-wide association signals: Application to hematological traits. *PLoS Genet*. 2011; 7(6):e1002139. [PubMed: 21738486]
43. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, Thorleifsson G, Helgadottir H, Steinthorsdottir V, Stefansson H, Williams C, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet*. 2009; 41(3):342–347. [PubMed: 19198610]
44. Ferreira MA, Mangino M, Brumme CJ, Zhao ZZ, Medland SE, Wright MJ, Nyholt DR, Gordon S, Campbell M, McEvoy BP, Henders A, et al. Quantitative trait loci for cd4:cd8 lymphocyte ratio are associated with risk of type 1 diabetes and hiv-1 immune control. *Am J Hum Genet*. 2010; 86(1):88–92. [PubMed: 20045101]
45. Okada Y, Kamatani Y, Takahashi A, Matsuda K, Hosono N, Ohmiya H, Daigo Y, Yamamoto K, Kubo M, Nakamura Y, Kamatani N. Common variations in psmd3-csf3 and plcb4 are associated with neutrophil count. *Hum Mol Genet*. 2010; 19(10):2079–2085. [PubMed: 20172861]
- 46•. Nalls MA, Couper DJ, Tanaka T, van Rooij FJ, Chen MH, Smith AV, Toniolo D, Zekai NA, Yang Q, Greinacher A, Wood AR, et al. Multiple loci are associated with white blood cell phenotypes. *PLoS Genet*. 2011; 7(6):e1002113. [PubMed: 21738480] [The three studies above use modern approaches to perform GWAS on WBC subsets and illustrate the value of using specific phenotypes to identify significant loci in different populations.]
- 47•. Reiner AP, Lettre G, Nalls MA, Ganesh SK, Mathias R, Austin MA, Dean E, Arepalli S, Britton A, Chen Z, Couper D, et al. Genome-wide association study of white blood cell count in 16,388 african americans: The continental origins and genetic epidemiology network (cogent). *PLoS Genet*. 2011; 7(6):e1002108. [PubMed: 21738479] [The three studies above use modern approaches to perform GWAS on WBC subsets and illustrate the value of using specific phenotypes to identify significant loci in different populations.]
- 48•. Okada Y, Hirota T, Kamatani Y, Takahashi A, Ohmiya H, Kumasaka N, Higasa K, Yamaguchi-Kabata Y, Hosono N, Nalls MA, Chen MH, et al. Identification of nine novel loci associated with white blood cell subtypes in a japanese population. *PLoS Genet*. 2011; 7(6):e1002067. [PubMed: 21738478] [The three studies above use modern approaches to perform GWAS on WBC subsets and illustrate the value of using specific phenotypes to identify significant loci in different populations.]

49. Maugeri N, Powell JE, t Hoen PA, de Geus EJ, Willemsen G, Kattenberg M, Henders AK, Wallace L, Penninx B, Hottenga JJ, Medland SE, et al. Lpar1 and itga4 regulate peripheral blood monocyte counts. *Hum Mutat.* 2011; 32(8):873–876. [PubMed: 21598361]
50. Anderlini P, Champlin RE. Biologic and molecular effects of granulocyte colony-stimulating factor in healthy individuals: Recent findings and current challenges. *Blood.* 2008; 111(4):1767–1772. [PubMed: 18057230]
51. Feral CC, Rose DM, Han J, Fox N, Silverman GJ, Kaushansky K, Ginsberg MH. Blocking the alpha 4 integrin-paxillin interaction selectively impairs mononuclear leukocyte recruitment to an inflammatory site. *J Clin Invest.* 2006; 116(3):715–723. [PubMed: 16470243]
52. Humbles AA, Lloyd CM, McMillan SJ, Friend DS, Xanthou G, McKenna EE, Ghiran S, Gerard NP, Yu C, Orkin SH, Gerard C. A critical role for eosinophils in allergic airways remodeling. *Science.* 2004; 305(5691):1776–1779. [PubMed: 15375268]
53. Gregersen PK, Diamond B, Plenge RM. Gwas implicates a role for quantitative immune traits and threshold effects in risk for human autoimmune disorders. *Curr Opin Immunol.* 2012; 24(5):538–543. [PubMed: 23026397]
54. Dendrou CA, Plagnol V, Fung E, Yang JH, Downes K, Cooper JD, Nutland S, Coleman G, Himsforth M, Hardy M, Burren O, et al. Cell-specific protein phenotypes for the autoimmune locus il2ra using a genotype selectable human bioresource. *Nat Genet.* 2009; 41(9):1011–1015. [PubMed: 19701192]
55. Sherborne AL, Hosking FJ, Prasad RB, Kumar R, Koehler R, Vijayakrishnan J, Papaemmanuil E, Bartram CR, Stanulla M, Schrappe M, Gast A, et al. Variation in cdkn2a at 9p21.3 influences childhood acute lymphoblastic leukemia risk. *Nat Genet.* 2010; 42(6):492–494. [PubMed: 20453839]
56. Sherborne AL, Houlston RS. What are genome-wide association studies telling us about b-cell tumor development? *Oncotarget.* 2010; 1(5):367–372. [PubMed: 21307401]
57. Crowther-Swanepoel D, Broderick P, Di Bernardo MC, Dobbins SE, Torres M, Mansouri M, Ruiz-Ponte C, Enjuanes A, Rosenquist R, Carracedo A, Jurlander J, et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. *Nat Genet.* 2010; 42(2):132–136. [PubMed: 20062064]
58. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, Kinsey SE, Lightfoot T, Roman E, Irving JA, Allan JM, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat Genet.* 2009; 41(9):1006–1010. [PubMed: 19684604]
59. Di Bernardo MC, Crowther-Swanepoel D, Broderick P, Webb E, Sellick G, Wild R, Sullivan K, Vijayakrishnan J, Wang Y, Pittman AM, Sunter NJ, et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nat Genet.* 2008; 40(10):1204–1210. [PubMed: 18758461]
60. Rossin EJ, Lage K, Raychaudhuri S, Xavier RJ, Tatar D, Benita Y, Cotsapas C, Daly MJ. Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. *PLoS Genet.* 2011; 7(1):e1001273. [PubMed: 21249183] [This article describes a valuable tool to examine protein-protein interactions among GWAS-implicated genes and discusses how this may provide insight into the biology underlying such traits.]
61. Auer PL, Johnsen JM, Johnson AD, Logsdon BA, Lange LA, Nalls MA, Zhang G, Franceschini N, Fox K, Lange EM, Rich SS, et al. Imputation of exome sequence variants into population-based samples and blood-cell-trait-associated loci in african americans: Nhlbi go exome sequencing project. *Am J Hum Genet.* 2012; 91(5):794–808. [PubMed: 23103231] [This recent study illustrates how rare variant studies of hematologic phenotypes can be conducted and will likely be the first of many other such studies that will be released in the future.]

Figure 1.

A simplified model of the hierarchy of hematopoiesis illustrates how GWAS of hematologic phenotypes can provide insight into the differentiation of various lineages. The diagram illustrates multipotent progenitors, including the long-term hematopoietic stem cells (LT-HSC), the short term HSC (ST-HSC), the common lymphoid progenitor (CLP), the common myeloid progenitor (CMP), the granulocyte-macrophage progenitor (GMP), and the megakaryocyte-erythroid progenitor (MEP). Committed progenitors leading to the various lineages are also shown. Phenotypes examined in the GWAS studies discussed in this review are shown at the bottom.