Physiol Genomics 50: 758–769, 2018.
First published June 29, 2018; doi:10.1152/physiolgenomics.00088.2017.

RESEARCH ARTICLE | *Genome-wide Association Studies and Function*

Identification of nine novel loci related to hematological traits in a Japanese population

Yoshiki Yasukochi,[1,2] Jun Sakuma,[2,3,4] Ichiro Takeuchi,[2,4,5] Kimihiko Kato,[1,6] Mitsutoshi Oguri,[1,7] Tetsuo Fujimaki,[8] Hideki Horibe,[9] and Yoshiji Yamada[1,2]

[1]*Department of Human Functional Genomics, Advanced Science Research Promotion Center, Mie University, Tsu, Mie, Japan;* [2]*CREST, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan;* [3]*Computer Science Department, College of Information Science, University of Tsukuba, Tsukuba, Ibaraki, Japan;* [4]*RIKEN Center for Advanced Intelligence Project, Tokyo, Japan;* [5]*Department of Computer Science, Nagoya Institute of Technology, Nagoya, Aichi, Japan;* [6]*Department of Internal Medicine, Meitoh Hospital, Nagoya, Aichi, Japan;* [7]*Department of Cardiology, Kasugai Municipal Hospital, Kasugai, Aichi, Japan;* [8]*Department of Cardiovascular Medicine, Inabe General Hospital, Inabe, Mie, Japan; and* [9]*Department of Cardiovascular Medicine, Gifu Prefectural Tajimi Hospital, Tajimi, Gifu, Japan*

Submitted 6 September 2017; accepted in final form 27 June 2018

Yasukochi Y, Sakuma J, Takeuchi I, Kato K, Oguri M, Fujimaki T, Horibe H, Yamada Y. Identification of nine novel loci related to hematological traits in a Japanese population. *Physiol Genomics* 50: 758–769, 2018. First published June 29, 2018; doi: 10.1152/physiolgenomics.00088.2017.—Recent genome-wide association studies have identified various genetic variants associated with hematological traits. Although it is possible that quantitative data of hematological traits are varied among the years examined, conventional genome-wide association studies have been conducted in a cross-sectional manner that measures traits at a single point in time. To address this issue, we have traced blood profiles in 4,884 Japanese individuals who underwent annual health check-ups for several years. In the present study, longitudinal exome-wide association studies were conducted to identify genetic variants related to 13 hematological phenotypes. The generalized estimating equation model showed that a total of 67 single nucleotide polymorphisms (SNPs) were significantly [false discovery rate (FDR) of <0.01] associated with hematological phenotypes. Of the 67 SNPs, nine SNPs were identified as novel hematological markers: rs4686683 of *SENP2* for red blood cell count (FDR = 0.008, $P = 5.5 \times 10^{-6}$); rs3917688 of *SELP* for mean corpuscular volume (FDR = 0.005, $P = 2.4 \times 10^{-6}$); rs3133745 of *C8orf37-AS1* for white blood cell count (FDR = 0.003, $P = 1.3 \times 10^{-6}$); rs13121954 at 4q31.2 for basophil count (FDR = 0.007, $P = 3.1 \times 10^{-5}$); rs7584099 at 2q22.3 (FDR = 2.6×10^{-5}, $P = 8.8 \times 10^{-8}$), rs1579219 of *HCG17* (FDR = 0.003, $P = 2.0 \times 10^{-5}$), and rs10757049 of *DENND4C* (FDR = 0.008, $P = 5.6 \times 10^{-5}$) for eosinophil count; rs12338 of *CTSB* for neutrophil count (FDR = 0.007, $P = 2.9 \times 10^{-5}$); and rs395967 of *OSMR-AS1* for monocyte count (FDR = 0.008, $P = 3.2 \times 10^{-5}$).

exome-wide association study; generalized estimating equation; hematological trait; linkage disequilibrium; longitudinal data

INTRODUCTION

Hematological traits are important for determining the health status or diagnosing diseases. Blood cells are classified into three major blood-cell lineages: red blood cells (RBCs), white blood cells (WBCs), and platelets. Recent genome-wide association studies (GWASs) have identified various genetic variants associated with hematological traits in diverse ethnic groups. Mousas et al. (2017) (29) examined relations of 137,086 rare coding or splice genetic variants [minor allele frequency (MAF) of <0.01] to 15 hematological traits in 308,572 subjects with European ancestry. They showed that 56 variants were associated with hematological phenotypes and identified a novel association of rs145535174 in *PLG* with platelet counts (29). Okada and Kamatani (2012) (31) systematically reviewed genetic association studies for hematological traits in several ethnic populations and reported that >100 genetic loci had been identified to affect the hematological phenotypes. Despite a number of associated loci, additional novel relations between genetic variants and hematological quantitative traits have been found to date (2, 16, 20, 35, 37). While some loci related to hematological traits are shared among ethnic populations, interethnic differences in the associated loci have been observed (15, 20, 23, 35). In Japan, however, there have been few GWASs for multiple hematological traits (15, 30). Hence, it is possible that genetic variants associated with hematological dynamics in Japanese individuals remain to be identified definitively.

Most conventional GWASs have been conducted in a cross-sectional manner that measured traits at a single point in time. Since the number of blood cells in Japanese adults is known to decrease with age (18), it is possible that quantitative data of hematological traits in an individual varied greatly among the years examined. However, the cross-sectional GWASs did not consider this possibility. Given that a longitudinal study examines longitudinal hematological data, this analysis increases the statistical power to detect the association. Therefore, hematological profiles in 4,884 Japanese individuals who underwent annual health check-ups for several years have been traced. We have now performed longitudinal exome-wide association studies to explore novel loci associated with 13 hematological traits: RBC count; hemoglobin (Hb); hematocrit (Ht); mean corpuscular volume (MCV); mean corpuscular hemoglobin (MCH); mean corpuscular hemoglobin concentration

Address for reprint requests and other correspondence: Y. Yasukochi, Dept. of Human Functional Genomics, Advanced Science Research Promotion Center, Organization for the Promotion of Regional Innovation, Mie Univ., 1577 Kurima-machiya, Tsu, Mie 514-8507, Japan (e-mail: hyasukou@proof.ocn.ne.jp).

  ISSN 1094-8341.

(MCHC); platelet count; WBC count; and neutrophil, basophil, eosinophil, lymphocyte, and monocyte counts.

MATERIALS AND METHODS

Compliance with ethical standards. The study protocol complied with the Declaration of Helsinki and was approved by the Committees on the Ethics of Human Research of Mie University Graduate School of Medicine and Inabe General Hospital. Written informed consent was obtained from all subjects before enrollment in the study.

Study subjects. Study subjects comprised a total of 4,884 community-dwelling individuals (2,792 men and 2,092 women) who visited the Health Care Center of Inabe General Hospital (Inabe, Mie, Japan) for annual health check-ups from April 2003 to March 2014. All participants had each undergone 1–11 medical examinations, and the average follow-up period, which was equivalent to the average number of measurements, was 5 yr. We refer to this cohort as the "Inabe cohort." Methods for the collection and storage of medical examination data and genomic DNA samples have been described previously (46). Patients with active cancers or pregnant women were not included in the study.

To assess genetic association, we used repeated measurements of 13 hematological traits in 4,884 Inabe subjects. The 13 traits consist of *1*) RBC count, the number of RBCs in peripheral venous blood (10^{12}/l); *2*) Hb, the concentration of hemoglobin in the blood (g/dl); *3*) Ht, the volume percentage of RBCs in the blood (%); *4*) MCV, the average volume of RBC that is expressed in femtoliters (fl), = Ht × 10/[RBC (10^{12}/l)]; *5*) MCH, the average mass of Hb per RBC in the sample (pg), = Hb × 10/[RBC (10^{12}/l)]; *6*) MCHC, the average concentration of Hb per RBC contained in the sample (g/dl), = Hb × 100/Ht; *7*) platelet count, the number of platelets in the blood (10^{9}/l); *8*) WBC count, the number of WBCs in the blood (10^{9}/l); and *9–13*) five WBC subtypes (neutrophils, basophils, eosinophils, lymphocytes, and monocytes), the percentage of each subtype in WBCs in the sample (%). We used an automated hematology analyzer LH-780 (Beckman Coulter, Brea, CA) to measure hematological traits of participants. The measurements were performed with a single equipment in the single clinical laboratory of the hospital.

The distributions of longitudinal data in the 13 hematological traits are shown in Fig. 1. In the longitudinal exome-wide association studies for RBCs, Hb, and MCH, a total of 4,884 subjects (2,792 men and 2,092 women) were examined (21,828–22,146 examinations). A total of 4,883 subjects (2,791 men and 2,092 women) were examined in the exome-wide association studies for Ht, MCV, and MCHC (21,828–22,133 examinations). Furthermore, 4,826 subjects (2,757 men and 2,069 women) and 4,833 subjects (2,762 men and 2,071 women) were examined in the exome-wide association studies for platelets (21,865 examinations) and WBCs (21,936 examinations), respectively. In the exome-wide association studies for the five WBC subtypes (neutrophils, basophils, eosinophils, lymphocytes, and monocytes), 1,073 subjects (5,591 examinations) for monocytes and 1,072 subjects (5,589–5,590 examinations) for the other subtypes were examined.

Longitudinal exome-wide association study. Longitudinal exome-wide association studies for the Inabe subjects were conducted, based

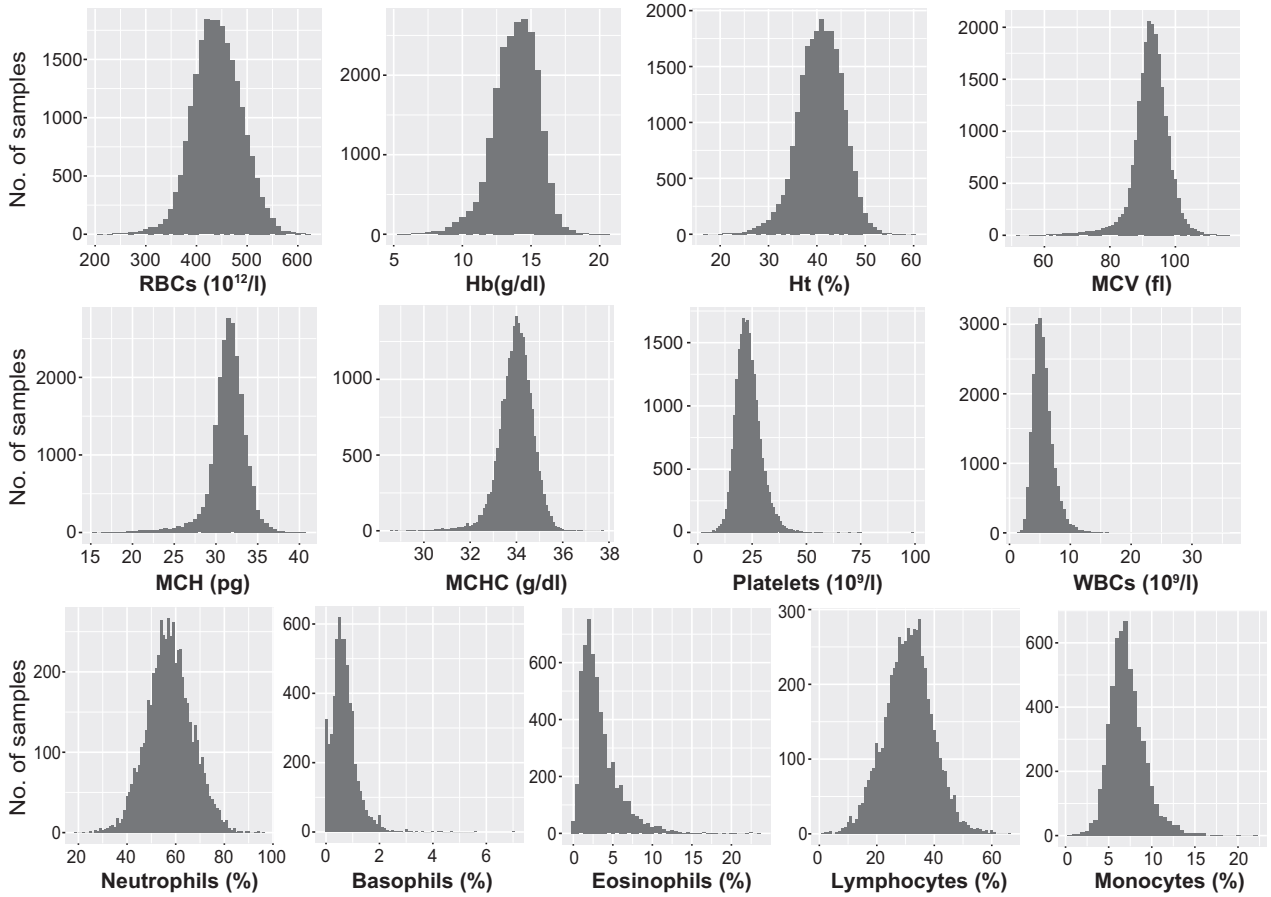

Fig. 1. Count distributions for longitudinal data of 13 hematological traits examined in the Inabe cohort (4,884 individuals). RBC, red blood cell count; Hb, hemoglobin; Ht, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; WBC, white blood cell count.

on ~244,000 genetic variants and longitudinal data of the 13 hematological traits. Genotyping for the subjects was performed with the Infinium HumanExome-12 ver. 1.2 BeadChip and Infinium Exome-24 ver. 1.0 BeadChip (Illumina, San Diego, CA). These arrays included putative functional exonic variants selected from >12,000 individual exome and whole-genome sequences across diverse ethnic populations, including European, African, Chinese, and Hispanic individuals (11). Approximately 3.6% of all the genetic variants were excluded from further analyses because one of the exome arrays did not contain them. We performed quality controls and discarded the following genotyped sites: *1*) monomorphic sites among Inabe subjects, *2*) variants with MAF of <0.05, *3*) variants whose genotype distribution significantly ($P < 0.001$) deviated from the Hardy-Weinberg equilibrium in the subjects. Analysis of the association of genetic variants on sex chromosomes with phenotypes is complicated because of the difference in the copy number between men and women and of X-inactivation in women. Analysis of the association of genetic variants on mitochondrial DNA with phenotypes is also complicated because of the existence of heteroplasmy. We thus discarded genetic variants located on sex chromosomes and mitochondrial DNA. Among the Inabe subjects, no individual was identified as population outliers by principal component analysis of genetic variants using the EIGENSTRAT method (33) via JMP Genomics version 6.0 (SAS Institute, Cary, NC) to detect population stratification on an exome-wide scale. Consequently, a total of 24,642 single nucleotide polymorphisms (SNPs) among 4,884 samples passed quality control.

Given that effects of SNPs on phenotypes may differ among three inheritance models, we examined the association of SNPs with hematological phenotypes in three inheritance models. We thus converted the genotyping data of 4,884 Inabe subjects into numeric data with dominant, additive, and recessive models. The dominant and recessive models were defined as "AA (0) vs. AB + BB (1)" and "AA + AB (0) vs. BB (1)" (A, major allele; B, minor allele), respectively, whereas the additive model was defined as "AA (0), AB (1), BB (2)." Quantile-quantile plots for *P* values of allele frequencies in the exome-wide association studies are shown in Figs. 2–4. The genomic inflation factor (λ) of *P* values was 1.00–1.14 for the 13 hematological traits in the dominant and additive models (Figs. 2 and 3). In the recessive model, the λ was 1.05–1.18 (Fig. 4). The rearrangement of Inabe longitudinal data was conducted in the R software version 3.32 (34) via RStudio version 1.0.136 (36) and by the use of a Perl script. We surveyed the linkage disequilibrium (LD) between the focal SNPs and other genetic variants that have been shown to be associated with hematological traits in previous studies through LDlink web-based tools (https://ldlink.nci.nih.gov) (25).

Statistical analyses. We examined relations between SNPs and repeated measurements of hematological traits by the generalized estimating equation (GEE) model (14, 21) with adjustments for age

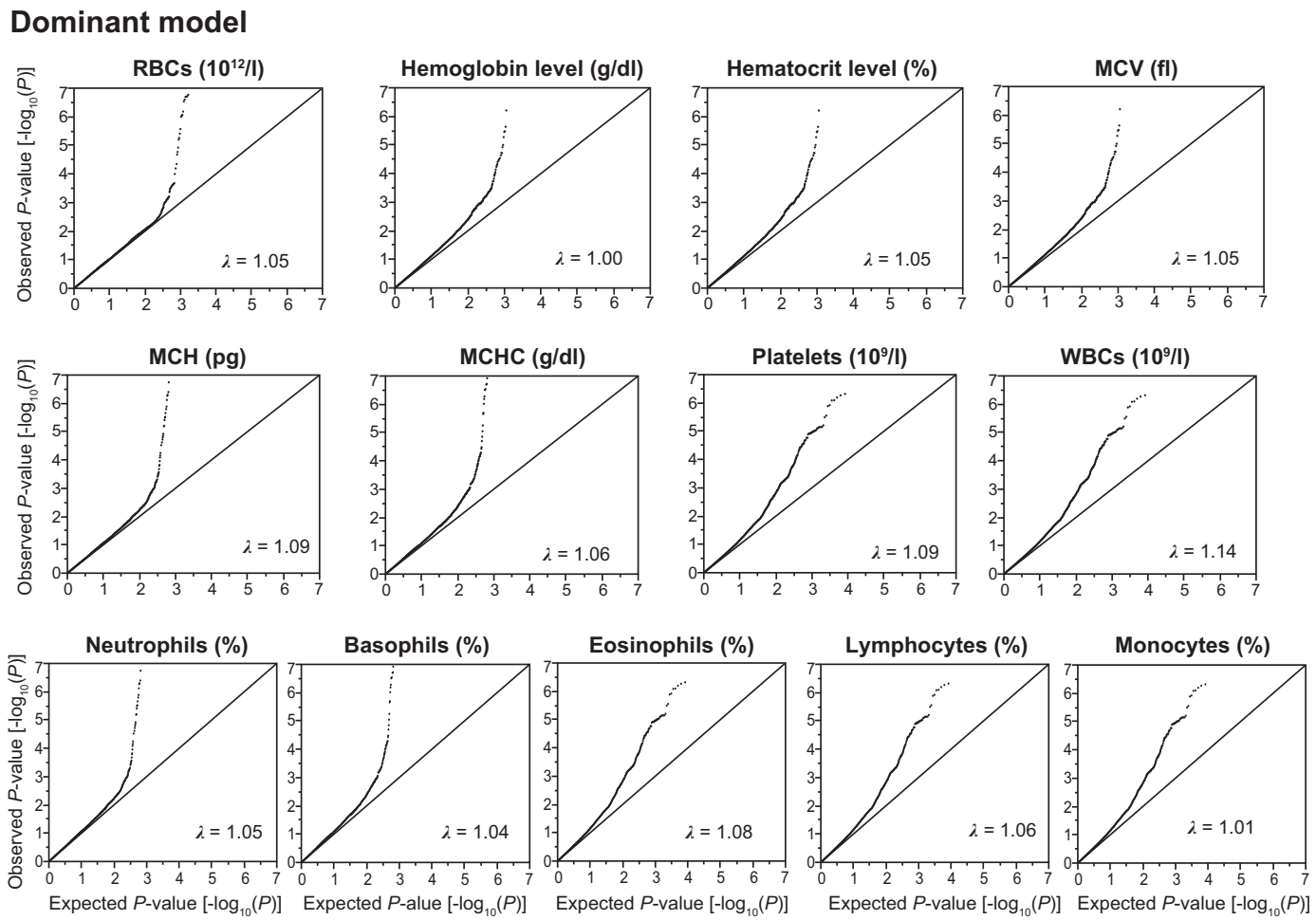

Fig. 2. Quantile-quantile plots for *P* values in the longitudinal exome-wide association studies for 13 hematological traits in the dominant model. The observed *P* values (*y*-axis) were compared with the expected *P* values (*x*-axis) under the null hypothesis, with the values being plotted as $-\log_{10}(P)$. RBC, red blood cell count; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; WBC, white blood cell count; λ represents the genomic inflation factor.

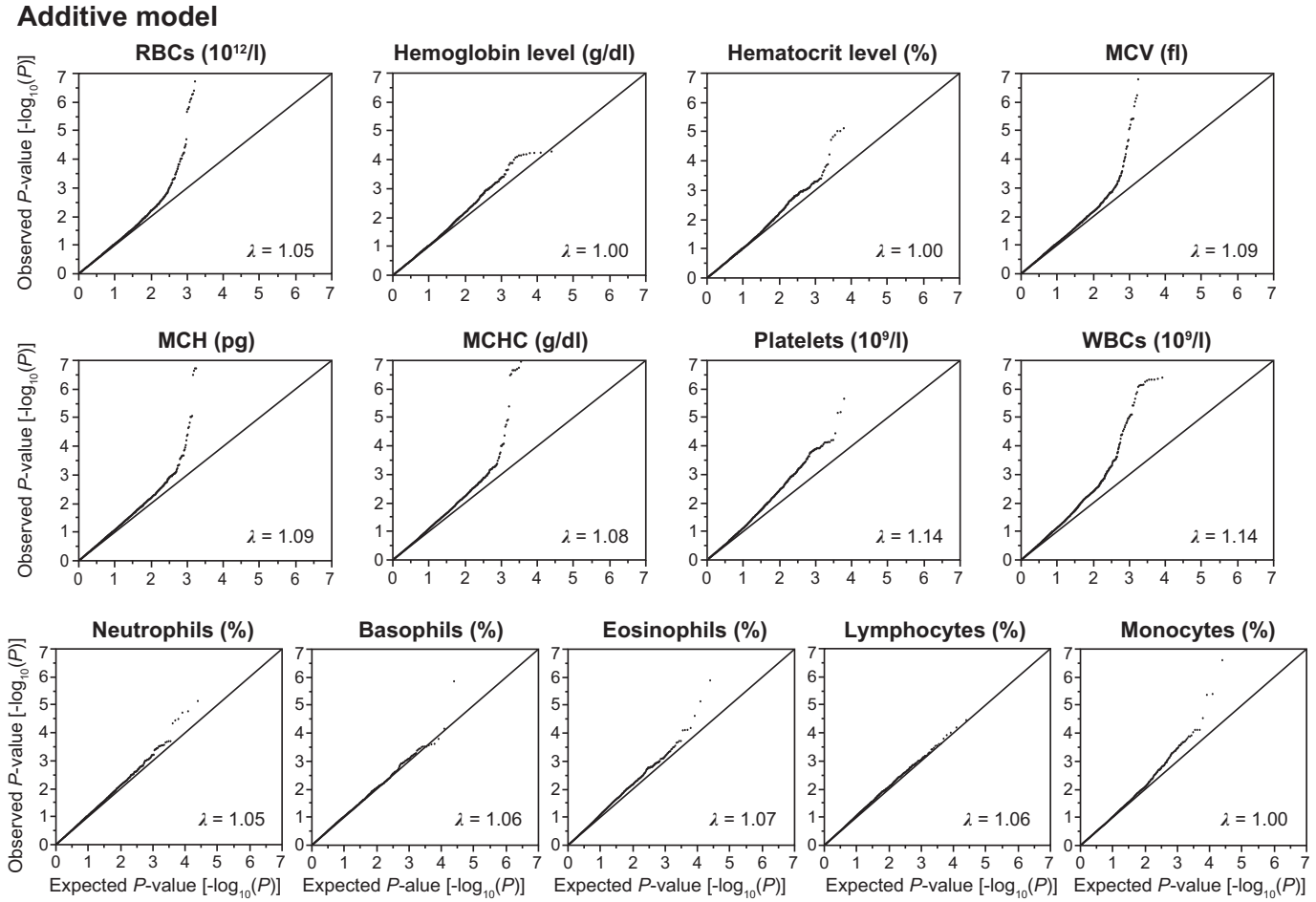

Fig. 3. Quantile-quantile plots for *P* values in the longitudinal exome-wide association studies for 13 hematological traits in the additive model. The observed *P* values (*y*-axis) were compared with the expected *P* values (*x*-axis) under the null hypothesis, with the values being plotted as $-\log_{10}(P)$. RBC, red blood cell count; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; WBC, white blood cell count. λ represents the genomic inflation factor.

and sex done with the R package "geepack" (13). The "waves" argument was used to specify the ordering of repeated measurements within individuals, because this argument is required to recognize clusters of each individual. Given that many SNPs in exome arrays are in LD, effects of SNPs on hematological traits are not independent. Therefore, we calculated the false discovery rate (FDR) by the Benjamin and Hochberg method (4) to compensate for multiple comparison of genotypes with the hematological parameters. An FDR of <0.01 was considered for statistical significance of association. Sitlani et al. (2015) (40) reported that a small effective sample size can increase the chances of generating type I errors. They recommended the use of "approxdf" for an effective sample size: approxdf $= 2 \times$ MAF $\times$ Nindep, where Nindep is the sum of the estimated number of independent observations per person, and an approxdf of ≥ 10 could reduce type I errors. Therefore, we estimated the approxdf for each candidate SNP by the R package "bosswithdf" (40, 44). To solve the issue of association with small effective sample sizes, we set to strict approxdf threshold, and discarded SNPs with approxdf of ≤ 30. The GEE model with the R package geepack was used to examine the association of SNPs with repeated measurements. To consider the association of SNPs with time-dependent phenotypic changes, we also calculated the *P* value with GEE using a *t* reference distribution with Satterthwaite estimates of degrees of freedom by the bosswithdf package, and incorporated a SNP $\times$ time interaction term into this model for time dependent analyses. The GEE with the *t* reference distribution was implemented using the longitudinal hematological data of subjects in the last five years (equivalent to the mean follow-up period).

Prediction of functional association for candidate loci. Gene-gene functional interactions were predicted using GeneMANIA Cytoscape plugin (27, 28, 45) via Cytoscape version 3.4.0 (38) software. Relations of candidate SNPs to hematological phenotypes reported by previous studies were investigated by GRASP (Genome-Wide Repository of Associations Between SNPs and Phenotypes, https://grasp.nhlbi.nih.gov/Overview.aspx) (19) and GWAS Catalogue (https://www.ebi.ac.uk/gwas/) (24) databases.

RESULTS

Characteristics of subjects. The characteristics of 4,884 Inabe subjects with respect to the longitudinal data of 13 hematological traits are shown in Table 1. The mean ages of men and women were 52.0 yr (range: 19–89 yr) and 52.2 yr (range: 18–91 yr of age), respectively.

Longitudinal exome-wide association studies for RBC traits and platelets. We assessed significant association between 13 hematological traits and SNPs that passed quality control in the three inheritance models with the GEE model. Candidate SNPs detected in all the inheritance models are

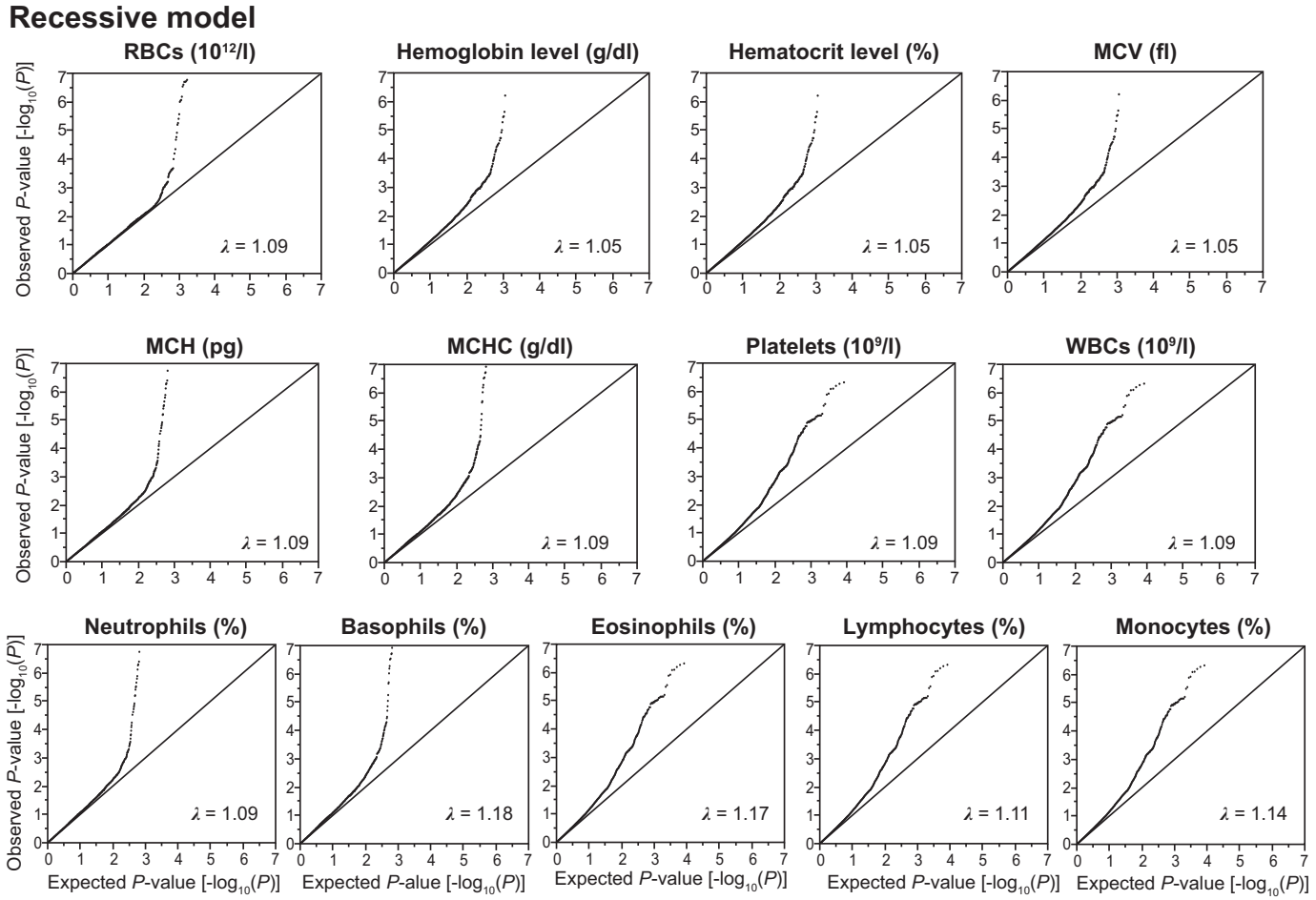

Fig. 4. Quantile-quantile plots for *P* values in the longitudinal exome-wide association studies for 13 hematological traits in the recessive model. The observed *P* values (*y*-axis) were compared with the expected *P* values (*x*-axis) under the null hypothesis, with the values being plotted as $-\log_{10}(P)$. RBC, red blood cell count; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; WBC, white blood cell count. λ represents the genomic inflation factor.

shown in Supplemental Table S1. (The online version of this article contains supplemental material.) Our longitudinal exome-wide association studies showed that a total of 67 SNPs were significantly (FDR <0.01) associated with hematological traits examined (Supplemental Table S2). Allele frequencies of these candidate SNPs are shown in Supplemental Table S3.

In longitudinal exome-wide association studies for six RBC traits, the GEE model showed 26 SNPs associated with RBCs, three SNPs with Ht, 25 SNPs with MCV, 17 SNPs with MCH, and 14 SNPs with MCHC, whereas there were no SNPs associated with Hb (Supplemental Tables S1 and S2). According to GRASP and GWAS Catalogue databases, 10 SNPs have not been reported by previous studies. However, previously known SNPs were located around the eight of 10 candidate SNPs, and these SNPs (four SNPs around *ABO* at the chromosomal region 9q34.2 and the different set of four SNPs around *ALDH2* at 12q24.1) were possibly in LD with the known SNPs related to the RBC traits examined. Thus, we did not consider candidate SNPs that were possibly correlated with the known associated SNPs as novel hematological markers. Consequently, we newly identified two SNPs related to RBC traits: rs4686683 of *SENP2* for RBCs (FDR = 0.008, $P = 5.5 \times 10^{-6}$, approxdf = 223) and rs3917688 of *SELP* for MCV (FDR = 0.005, $P = 2.4 \times 10^{-6}$, approxdf = 45) (Supplemental Table S2). Mean quantitative values of each hematological measurement in individuals with different genotypes at the novel candidate SNPs are shown in Supplemental Table S4.

In the longitudinal exome-wide association studies for platelets, two SNPs were significantly related to platelet count (Supplemental Tables S1 and S2). However, these SNPs have been shown to be associated with hematological phenotypes, according to GRASP and GWAS Catalogue databases. Therefore, there were no novel genetic determinants for platelets in our longitudinal exome-wide association studies.

Longitudinal exome-wide association study for WBC traits. We performed longitudinal exome-wide association studies for six WBC traits and found 19 SNPs associated with WBCs, one SNP with basophils, three SNPs with eosinophils, one SNP with neutrophils, and two SNPs with monocytes (Supplemental Table S2). No significant association with lymphocytes was observed in the present study. Of the 26 SNPs, the association of 24 has not been reported by previous studies according to the two databases. However, 17 were possibly in LD with previously reported SNPs related to WBC traits.

Table 1. *Longitudinal characteristics of 4,884 Inabe subjects*

Characteristic	Subjects, *n*	Examinations, *n*	Means ± SD	Range
Age, yr	4,884	22,146	52.1 ± 12.8	18–91
RBCs, 10^{12}/l	4,884	22,145	4.4 ± 0.5	2.1–6.2
Hb, g/dl	4,884	22,146	13.9 ± 1.6	5.5–20.6
Ht, %	4,883	22,133	40.8 ± 4.6	17.4–59.9
MCV, fl	4,883	21,828	92.5 ± 5.6	52.1–115.9
MCH, pg	4,884	21,828	31.5 ± 2.2	15.3–40.5
MCHC, g/dl	4,883	21,828	34.0 ± 0.7	28.6–37.7
Platelets, 10^{9}/l	4,826	21,865	234 ± 58	17–990
WBCs, 10^{9}/l	4,833	21,936	5.5 ± 1.7	1.5–36.1
Neutrophils, %	1,072	5,590	57.6 ± 9.3	18.5–96.4
Basophils, %	1,072	5,590	0.7 ± 0.5	0–7.0
Eosinophils, %	1,072	5,589	3.4 ± 2.5	0–23.5
Lymphocytes, %	1,072	5,590	31.1 ± 8.4	1.1–66.2
Monocytes, %)	1,073	5,591	7.2 ± 2.1	0.7–22.0

The number of subjects is based on data examined in the latest year. Examination values indicate the numbers of measurements taken. Quantitative data are means ± SD. RBC, red blood cell count; Hb, hemoglobin; Ht, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; WBC, white blood cell count.

Of the 17 SNPs, the GEE model showed significant relations of 15 SNPs across a ~569 kb genomic region at 6p21.33 (in the *HLA* region) to WBCs (Supplemental Tables S1 and S2). Although there have been no reports of association between these SNPs and hematological traits examined, previous studies have reported association of SNPs in the *HLA* region with WBCs (15, 30, 31). On the basis of SNPs in the Inabe cohort, we generated an LD plot in a ~1.4 Mb genomic region at 6p21.3–p22.1 with the Haploview program version 4.2 (3) (Fig. 5). The plot showed moderate or strong correlations between pairs of SNPs examined. Since long-range haplotypes in the *HLA* region have been maintained under natural selection (12, 32), it is difficult to distinguish which SNP is responsible for hematological dynamics in individuals. Therefore, we did not consider the 15 SNPs at 6p21.33 as novel candidate hematological markers, although it is possible that any of these SNPs are actually associated with hematological traits.

The GEE model showed novel relations of rs56030650 in *GSDMA* and rs10107630 in *CCDC26* to WBCs and monocytes, respectively (Supplemental Table S2). However, other SNPs in these genes (rs3859192 in *GSDMA* and rs10956483 in *CCDC26*) have been shown to be related to the corresponding WBC traits (6, 30). LDpair, an LDlink application, indicated that rs56030650 and rs10107630 were in LD with rs3859192 ($r^2 = 0.98$, $D' = 1.00$) and rs10956483 ($r^2 = 0.88$, $D' = 0.94$), respectively, in JPT (Japanese in Tokyo, Japan) from the 1000 Genomes Project database (http://www.internationalgenome.org) (43). Hence, we did not consider the two candidate SNPs as novel hematological markers.

We thus identified the following seven SNPs as novel genetic determinants for WBC traits: rs3133745 of *C8orf37-AS1* for WBCs (FDR = 0.003, $P = 1.3 \times 10^{-6}$, approxdf = 163), rs13121954 at the chromosomal region 4q31.2 for basophils (FDR = 0.007, $P = 3.1 \times 10^{-5}$, approxdf = 47), rs7584099 at 2q22.3 for eosinophils (FDR = 2.6×10^{-5}, $P = 8.8 \times 10^{-8}$, approxdf = 33), rs1579219 of *HCG17* for eosinophils (FDR = 0.003, $P = 2.0 \times 10^{-5}$, approxdf = 37), rs10757049 of *DENND4C* for eosinophils (FDR = 0.008, $P = 5.6 \times 10^{-5}$, approxdf = 40), rs12338 of *CTSB* for neutrophils (FDR = 0.007, $P = 2.9 \times 10^{-5}$, approxdf = 49), and rs395967 of *OSMR-AS1* for monocytes (FDR = 0.008, $P = 3.2 \times 10^{-5}$, approxdf = 49).

Additional exome-wide association studies. To examine the association of the nine candidate SNPs with the time-varying phenotypes, we applied the GEE model with a *t* reference distribution including a SNP × time interaction term (see MATERIALS AND METHODS). This analysis showed that four of nine candidate SNPs were associated ($P < 0.05$) with the longitudinal changes of hematological trait (rs4686683 of *SENP2* for RBCs, rs13121954 at 4q31.2 for basophils, rs12338 of *CTSB* for neutrophils, and rs1579219 of *HCG17* for eosinophils).

In the present study, an FDR of <0.01 is considered statistically significant. If Bonferroni's correction [a *P* value of $<2.0 \times 10^{-6}$ (0.05/24,642 SNPs)] is applied, relations of rs7584099 at 2q22.3 ($P = 8.8 \times 10^{-8}$) and rs3133745 of *C8orf37-AS1* ($P = 1.3 \times 10^{-6}$) to eosinophils and WBCs, respectively, were significant. Since the remaining seven SNPs did not reach the threshold of *P* value with Bonferroni's correction, the association of the seven SNPs with hematological phenotypes should be carefully interpreted, and replication studies are required to verify the association.

To detect the association of rare population-specific variants, we conducted longitudinal exome-wide association studies for hematological traits, using 41,442 genetic variants including low-frequency and rare variants ($0.1 \leq MAF \leq 5\%$). Even though a large number of genetic variants were significantly (FDR <0.01) associated with hematological phenotypes, all of these variants with a MAF of <5% were filtered out by an approxdf threshold of >30.

Association of newly identified and previously reported SNPs with hematological phenotypes. We examined 67 SNPs related to hematological phenotypes in the present study with information of *P* values in three Blood-Cell Consortium (BCX) studies (5, 9, 41) from the website Laboratory of Guillaume Lettre (http://www.mhi-humangenetics.org). All the three, 17, two, and one SNPs associated with Ht, MCH, platelets, and basophils, respectively, in our longitudinal exome-wide association studies were also related ($P < 0.05$) to their phenotypes in the previous studies (Supplemental Table S5). There were no SNPs associated with eosinophils in the previous studies, although three SNPs were related to eosinophils in our longitudinal exome-wide association studies. The association of many SNPs was observed in the previous studies: nine of 14 SNPs associated with MCHC, 19 of 23 SNPs associated with MCV, one of two SNPs associated with monocytes, 22 of 25 SNPs associated with RBC counts, and 13 of 18 SNPs associated with WBC counts. Of the nine SNPs that were identified as novel hematological markers in the present study, rs13121954 were associated with basophils in both previous and present studies. In the BCX studies, no information on *P* values for rs12338 was available, and six SNPs were related to different hematological phenotypes from the present study: rs3917688 associated with basophils and platelets, rs395967 with red blood cell distribution width (RDW), rs1579219 with neutrophils, rs3133745 with RDW and neutrophils, and rs10757049 with MCH and RBC. Two SNPs (rs7584099 and rs4686683) were not associated with any hematological phenotypes in the previous studies.

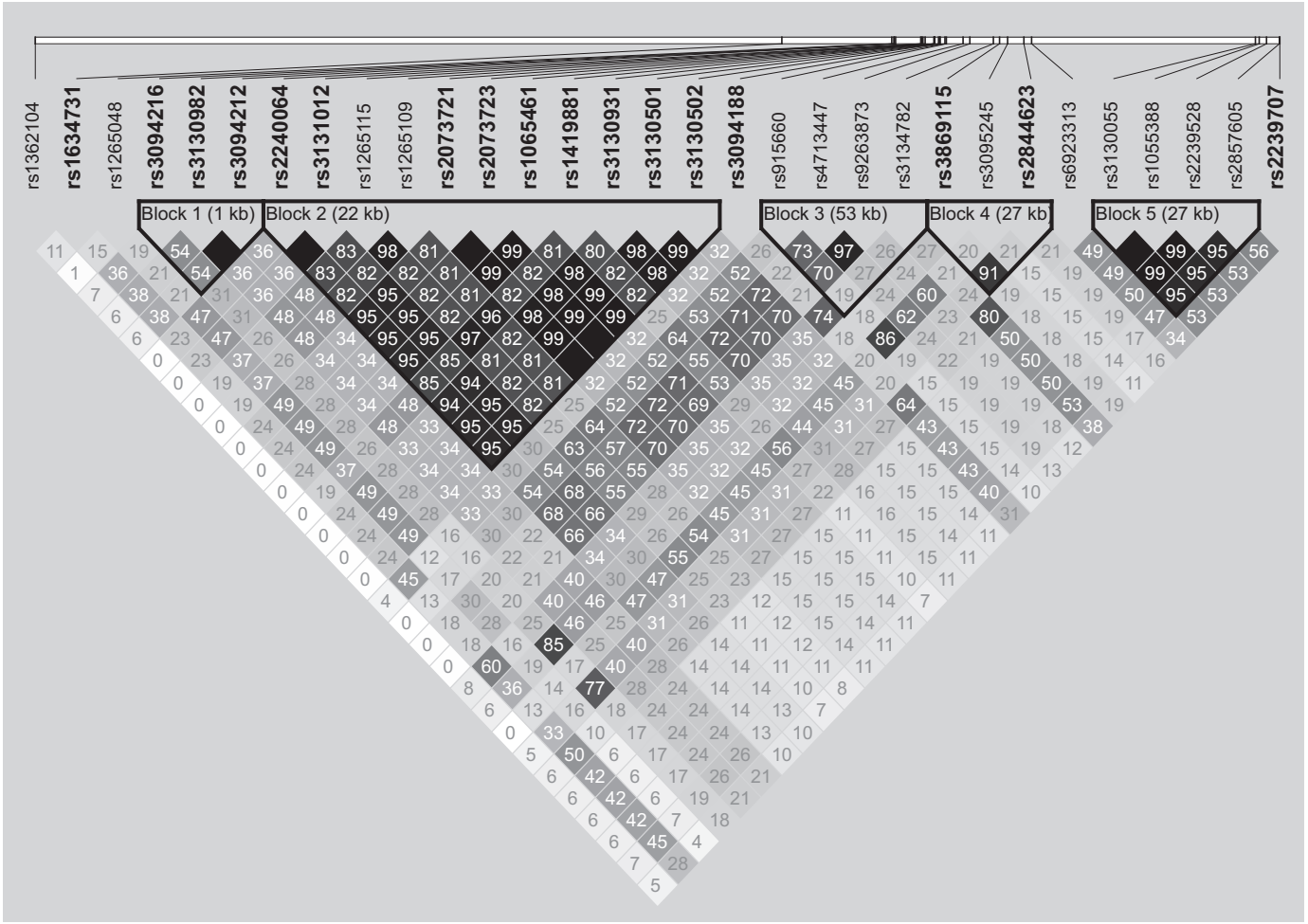

Fig. 5. Linkage disequilibrium of biallelic sites in a ~1.4 Mb genomic region at 6p21.3–p22.1 based on single nucleotide polymorphism (SNP) data from 4,884 Inabe residents. The SNPs with the minor allele frequency of <0.01 were removed from the analysis. The white blood cell-associated SNPs are shown in boldface.

We compared the effect direction and allele frequency of effect allele in subjects with East Asian and European ancestries (Supplemental Table S6). In the comparison of effect direction of 63 SNPs between Inabe subjects (East Asian ancestry) and Europeans in BCX studies, the consistency of the direction in 20–48 SNPs was observed: 34 SNPs for basophils, 36 for eosinophils, 48 for Ht, 45 for Hb, 28 for lymphocytes, 44 for MCH, 39 for MCHC, 43 for MCV, 20 for monocytes, 29 for neutrophils, 47 platelets, 41 for RBC counts, and 40 for WBC counts. Of the eight newly identified SNPs (no information of *P* values for rs12338), three (rs7584099, rs4686683, and rs1579219) showed the same direction of association. Although frequencies of effect alleles were different between the subjects with East Asian and European ancestries, the effect direction was the same [e.g., the effect direction of association in rs4686683 was same between the two groups although the frequencies of effect allele were different between Inabe subjects (0.399) and Europeans (0.613)].

We also examined 494 variants (including variants with an MAF of <0.05) identified in the BCX studies ($P < 2.0 \times 10^{-7}$, significance level in the studies), using data of our longitudinal exome-wide association studies (Supplemental Table S7). Of the 494 variants, 329 were associated ($P < 0.05$) with at least one of hematological phenotypes examined in the additive model. Of these variants examined, moreover, 41–92 SNPs were associated with each hematological trait examined: 90 SNPs for RBC counts, 86 for Hb, 83 for Ht, 83 for MCV, 91 for MCH, 89 for MCHC, 92 for platelets, 79 for WBC counts, 47 for neutrophils, 38 for basophils, 42 for eosinophils, 41 for lymphocytes, and 55 for monocytes.

We additionally examined the association of 67 SNPs with hematological traits examined in our exome-wide association studies, using the data set of Astle et al. (2016) (1) in the GRASP database (Supplemental Table S8). The information on *P* values for 40 SNPs including seven newly identified SNPs is available in the data set. All SNPs associated with Ht (3 SNPs), MCH (16), MCHC (8), platelets (2), and eosinophils (1) in our longitudinal exome-wide association studies were also related ($P < 0.05$) to the corresponding phenotypes in the previous studies. The rs13121954 was not associated with basophils. The number of difference of SNPs associated with the following phenotypes was only one between present and previous studies: MCV (18 SNPs), monocytes (1), RBC counts (19), and WBC counts (6) in the previous study. Of the seven newly identified SNPs, rs7584099 and rs4686683 were associated with eosinophils and RBCs, respectively, in both previous and present studies.

In the comparison of effect direction of 40 SNPs between the Inabe subjects and Europeans in Astle et al. (2016) (1), the consistency of the direction in 12–35 SNPs was observed: 24 SNPs for basophils, 26 for eosinophils, 33 for Ht, 35 for Hb, 16 for lymphocytes, 31 for MCH, 29 for MCHC, 31 for MCV, 12 for monocytes, 20 for neutrophils, 32 platelets, 32 for RBC counts, and 31 for WBC counts (Supplemental Table S9). Of the seven newly identified SNPs (no information on *P* values for rs12338 and rs1579219), four (rs3917688, rs4686683, rs395967, and rs3133745) showed the same direction of association.

We also examined 80 variants identified ($P < 2.0 \times 10^{-7}$) in Astle et al. (2016) (1), by the use of data in our longitudinal exome-wide association studies (Supplemental Table S10) and observed that 50 variants were related ($P < 0.05$) to at least one of hematological phenotypes in the additive model. Of the 80 variants examined, 7–13 SNPs were associated with each hematological trait examined: 11 SNPs for RBC counts, seven for Hb, nine for Ht, 11 for MCV, 10 for MCH, 11 for MCHC, 13 for platelets, seven for WBC counts, nine for neutrophils, eight for basophils, eight for eosinophils, eight for lymphocytes, and nine for monocytes.

Gene interaction network analysis. To explore interactive functional association among newly identified and previously known genes related to hematological dynamics, we preliminary performed a GeneMANIA network analysis (Fig. 6). The network displayed many potential direct or indirect association. The network analysis suggested that the DENN domain containing the 4C (*DENND4C*) gene interacts with the hemoglobin subunit alpha 1 (*HBA1*) and HBS1-like translational GTPase (*HBS1L*) genes (22). The SUMO1/sentrin/SMT3-specific peptidase 2 (*SENP2*) gene was coexpressed with the egl-9 family hypoxia-inducible factor 1 (*EGLN1*) gene and was indirectly associated with many genes related to hematological traits via selenoprotein T (*SELENOT*) and small ubiquitin-like modifier 1 (*SUMO1*). Cathepsin B (*CTSB*) and oncostatin M receptor (*OSMR*) genes appeared to be directly associated with many genes related to hematological traits. The selectin P (*SELP*) gene showed shared protein domains and colocalization with the glycoprotein Ib platelet alpha subunit (*GP1BA*) gene. There was no information on association with the chromosome 8 open reading frame 37 (*C8orf37*) gene and HLA complex group 17 (*HCG17*) noncoding RNA in the analysis.

DISCUSSION

In the longitudinal exome-wide association studies for 13 hematological traits, a total of 67 SNPs were significantly

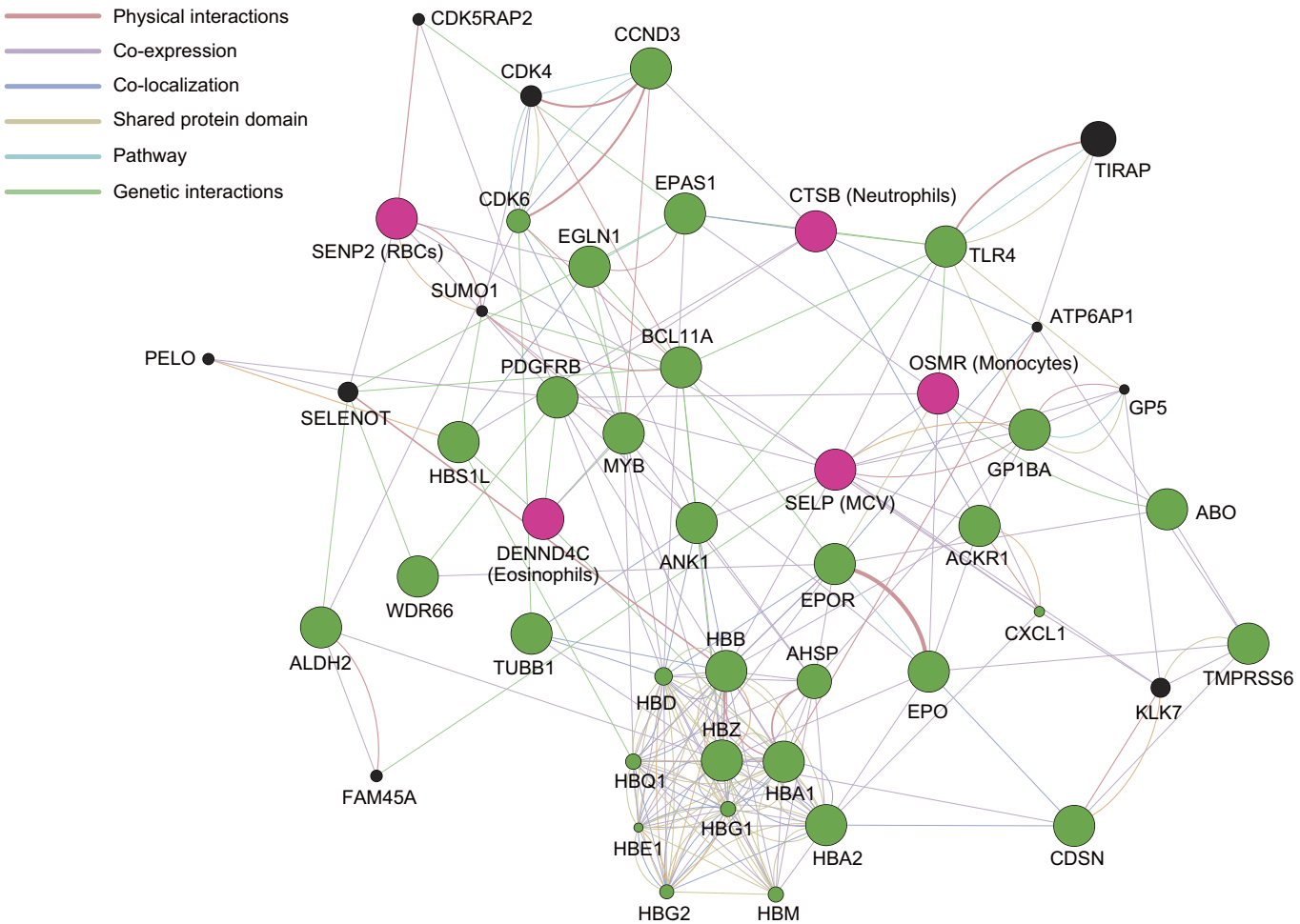

Fig. 6. Diagram of GeneMANIA network analysis for newly identified (pink circles) and previously reported (green circles) genes related to hematological dynamics. The hematological traits associated with newly identified SNPs are shown in parentheses.

(FDR <0.01) related to hematological traits. Of the 67 SNPs, nine (rs4686683 of *SENP2* for RBCs; rs3133745 of *C8orf37-AS1* for WBCs; rs3917688 of *SELP* for MCV; rs13121954 at 4q31.2 for basophils; rs12338 of *CTSB* for neutrophils; rs395967 of *OSMR-AS1* for monocytes; and rs7584099 at 2q22.3, rs1579219 of *HCG17*, and rs10757049 of *DENND4C* for eosinophils) were identified as novel hematological markers (Supplemental Tables S1 and S2). The MAFs of the nine newly identified and 58 previously reported SNPs were 24–49% and 12–50%, respectively, and approxdf values of those SNPs were 33–223 and 35–647, respectively. Our longitudinal exome-wide association studies thus identified common variants related to hematological traits.

The GeneMANIA network suggested some potential direct or indirect relations between genes associated with hematological phenotypes and newly identified genes. In addition, some of the nine loci have plausible function. SENP2 protein is involved in the process of synthesized SUMO1 into conjugate form and activates to remove SUMO1 from its substrate. The SUMO protein plays an important role in essential biological function. This protein may cause hematological diseases such as acute myeloid leukemia (8, 47). Since the SENP2 protein is involved in the regulation of SUMOylation, the nucleotide substitution (C→A) at rs4686683 may affect the hematological dynamics. However, the candidate SNP is an intron variant. Therefore, the functional relevance of the candidate SNP to the hematological dynamics remains unclear.

The nonsynonymous substitution (G→C, L26V) at rs12338 in *CTSB* was significantly related to neutrophils in our longitudinal exome-wide association studies. *CTSB* is ubiquitously expressed in various tissues and organs, according to The Human Protein Atlas (http://www.proteinatlas.org/). This gene is a member of the C1 family of peptidases. The CTSB enzyme is released by neutrophils and cleaves the extracellular domain in CD18 molecules that play important roles in the regulation of neutrophil adhesion and migration (10, 39). Given that the CTSB enzyme is involved in neutrophil adhesion and migration through the CD18 cleavage, the association of *CTSB* with neutrophils may be attributable to the effect of this enzyme on the regulation in neutrophil dynamics.

The rs395967 (A→G) of *OSMR-AS1* showed significant association with monocytes. OSMR antisense RNA 1 (*OSMR-AS1*) is a noncoding RNA. *OSMR* is a member of interleukin (IL)-6 family cytokine receptors and is expressed in various tissues and organs (The Human Protein Atlas). Expression of *OSMR* and interleukin 31 receptor A (*IL31RA*) mRNA is induced in activated monocytes, and *OSMR* plays a role in inflammatory reactions and hematopoiesis (7, 42, 48). Given that *OSMR-AS1* regulates the expression of *OSMR*, the association of *OSMR-AS1* with monocytes might be attributable to the effect of this noncoding RNA on the regulation in monocyte dynamics.

The SELP protein is a 140 kDa adhesion molecule involved in the interaction of platelets with leukocytes. This protein is mainly distributed in megakaryocytes and platelets as well as vascular endothelial cells of diverse human organs (26). Knock-in mice with elevated plasma levels of soluble SELP showed several abnormalities related to atherosclerosis and cerebrovascular complications (17). In our longitudinal exome-wide association studies, rs3917688 (G→A) in *SELP* was significantly associated with MCV. However, this SNP was not associated with MCV in BCX studies (5, 9, 41) and Astle et al. (2016) (1) (P = 0.205–0.781). The minor allele "A" of rs3917688 in Asians (0.329–0.372) was major in European (0.509) and African populations (0.577), according to the 1000 Genomes Project. The difference of allele frequencies might be attributable to the discrepancy of association between SNPs and hematological phenotypes among ethnic groups. Although the effect direction of association in rs3917688 was same in the present study and the study of Astle et al. (2016), the effect of this SNP appeared to be stronger in Inabe subjects with East Asian ancestry (0.2748) than in subjects with European ancestry (0.0045). The functional relevance of SELP protein to MCV remains unclear, and further functional analysis is thus required to clarify the results of this study.

In the present study, SNPs in two noncoding RNA (*HCG17* and *C8orf37-AS1*) and one coding gene (*DENND4C*) were related to hematological traits. However, these noncoding RNA and gene have not been reported to show any association with hematological traits to date. Therefore, further functional analysis is required to verify the relations of these SNPs to hematological phenotypes. Of the nine SNPs newly identified in the present study, seven (rs7584099, rs3917688, rs13121954, rs12338, rs395967, rs1579219, and rs10757049) were detected from the GEE method in the recessive model only. It is possible that the recessive model may increase the chances of generating type I error due to potential genomic inflations. Therefore, replication studies are required to verify the association of the seven SNPs with hematological traits.

There are several limitations in the presents study. First, the longitudinal exome-wide association studies were conducted in a local Japanese population. Seven of the nine identified SNPs did not reach the threshold of P value with Bonferroni's correction. In the examinations of WBC subtypes, the small sample size decreased the statistical power. Therefore, replication studies in other Japanese populations or other ethnic groups are required to verify the association of identified SNPs with hematological phenotypes. Second, the functional relevance of the nine SNPs identified in the present study to hematological traits remains unclear. Further functional analysis is required to clarify the results of this study. Third, genetic variants on sex chromosomes and mitochondrial DNA were not examined in our longitudinal exome-wide association studies. Forth, the follow-up period of annual health check-ups varied from 1 to 11 yr among individuals. An exome-wide association scan that tests the association of SNPs with time-dependent phenotypic changes is important. However, the time-dependent analysis with the R package bosswithdf does not allow any missing data. In future, a genome-wide time-dependent analysis is required to elucidate the association of SNPs with the time-varying phenotypes.

In conclusion, the GEE model showed that nine SNPs (rs4686683 of *SENP2*, rs3133745 of *C8orf37-AS1*, rs3917688 of *SELP*, rs13121954 at 4q31.2, rs12338 of *CTSB*, rs395967 of *OSMR-AS1*, rs7584099 at 2q22.3, rs1579219 of *HCG17*, and rs10757049 of *DENND4C*) were identified as novel hematological markers. The results of our longitudinal exome-wide association studies may contribute to the hematological study, and genotyping for these SNPs may be useful for precision medicine in Japanese.

ACKNOWLEDGMENTS

The authors deeply thank all the subjects for their participation in the study.

GRANTS

This work was supported by Research Grant from Okasan Kato Culture Promotion Foundation (to Y. Yasukochi), the Kurata Grant (no. 1323) awarded by the Hitachi Global Foundation (to Y. Yasukochi, Y. Yamada), CREST (JPMRJCR1302) of the Japan Science and Technology Agency (to Y. Yamada, J. Sakuma, I. Takeuchi), and by Japan Society for the Promotion of Science KAKENHI Grants JP17H00758 (to I. Takeuchi, Y. Yasukochi) and JP15H04772 (to Y. Yamada).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

Y. Yasukochi and Y. Yamada conceived and designed research; Y. Yasukochi, J.S., I.T., and Y. Yamada analyzed data; Y. Yasukochi, J.S., I.T., and Y. Yamada interpreted results of experiments; Y. Yasukochi prepared figures; Y. Yasukochi drafted manuscript; Y. Yasukochi, J.S., I.T., K.K., M.O., T.F., H.H., and Y. Yamada edited and revised manuscript; Y. Yasukochi, J.S., I.T., K.K., M.O., T.F., H.H., and Y. Yamada approved final version of manuscript; K.K., M.O., T.F., H.H., and Y. Yamada performed experiments.

REFERENCES

1. **Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, Lambourne JJ, Sivapalaratnam S, Downes K, Kundu K, Bomba L, Berentsen K, Bradley JR, Daugherty LC, Delaneau O, Freson K, Garner SF, Grassi L, Guerrero J, Haimel M, Janssen-Megens EM, Kaan A, Kamat M, Kim B, Mandoli A, Marchini J, Martens JHA, Meacham S, Megy K, O'Connell J, Petersen R, Sharifi N, Sheard SM, Staley JR, Tuna S, van der Ent M, Walter K, Wang SY, Wheeler E, Wilder SP, Iotchkova V, Moore C, Sambrook J, Stunnenberg HG, Di Angelantonio E, Kaptoge S, Kuijpers TW, Carrillo-de-Santa-Pau E, Juan D, Rico D, Valencia A, Chen L, Ge B, Vasquez L, Kwan T, Garrido-Martín D, Watt S, Yang Y, Guigo R, Beck S, Paul DS, Pastinen T, Bujold D, Bourque G, Frontini M, Danesh J, Roberts DJ, Ouwehand WH, Butterworth AS, Soranzo N.** The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* 167: 1415–1429.e19, 2016. doi:10.1016/j.cell.2016.10.042.
2. **Auer PL, Teumer A, Schick U, O'Shaughnessy A, Lo KS, Chami N, Carlson C, de Denus S, Dubé M-P, Haessler J, Jackson RD, Kooperberg C, Perreault L-PL, Nauck M, Peters U, Rioux JD, Schmidt F, Turcot V, Völker U, Völzke H, Greinacher A, Hsu L, Tardif J-C, Diaz GA, Reiner AP, Lettre G.** Rare and low-frequency coding variants in *CXCR2* and other genes are associated with hematological traits. *Nat Genet* 46: 629–634, 2014. doi:10.1038/ng.2962.
3. **Barrett JC, Fry B, Maller J, Daly MJ.** Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263–265, 2005. doi:10.1093/bioinformatics/bth457.
4. **Benjamini Y, Hochberg Y.** Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* 57: 289–300, 1995.
5. **Chami N, Chen MH, Slater AJ, Eicher JD, Evangelou E, Tajuddin SM, Love-Gregory L, Kacprowski T, Schick UM, Nomura A, Giri A, Lessard S, Brody JA, Schurmann C, Pankratz N, Yanek LR, Manichaikul A, Pazoki R, Mihailov E, Hill WD, Raffield LM, Burt A, Bartz TM, Becker DM, Becker LC, Boerwinkle E, Bork-Jensen J, Bottinger EP, O'Donoghue ML, Crosslin DR, de Denus S, Dubé MP, Elliott P, Engström G, Evans MK, Floyd JS, Fornage M, Gao H, Greinacher A, Gudnason V, Hansen T, Harris TB, Hayward C, Hernesniemi J, Highland HM, Hirschhorn JN, Hofman A, Irvin MR, Kähönen M, Lange E, Launer LJ, Lehtimäki T, Li J, Liewald DCM, Linneberg A, Liu Y, Lu Y, Lyytikäinen LP, Mägi R, Mathias RA, Melander O, Metspalu A, Mononen N, Nalls MA, Nickerson DA, Nikus K, O'Donnell CJ, Orho-Melander M, Pedersen O, Petersmann A, Polfus L, Psaty BM, Raitakari OT, Raitoharju E, Richard M, Rice KM, Rivadeneira F, Rotter JI, Schmidt F, Smith AV, Starr JM, Taylor KD, Teumer A, Thuesen BH, Torstenson ES, Tracy RP, Tzoulaki I, Zakai NA, Vacchi-Suzzi C, van Duijn CM, van Rooij FJ, Cushman M, Deary IJ, Velez Edwards DR, Vergnaud AC, Wallentin L, Waterworth DM, White HD, Wilson JG, Zonderman AB, Kathiresan S, Grarup N, Esko T, Loos RJ, Lange LA, Faraday N, Abumrad NA, Edwards TL, Ganesh SK, Auer PL, Johnson AD, Reiner AP, Lettre G.** Exome genotyping identifies pleiotropic variants associated with red blood cell traits. *Am J Hum Genet* 99: 8–21, 2016. doi:10.1016/j.ajhg.2016.05.007.
6. **Crosslin DR, McDavid A, Weston N, Nelson SC, Zheng X, Hart E, de Andrade M, Kullo IJ, McCarty CA, Doheny KF, Pugh E, Kho A, Hayes MG, Pretel S, Saip A, Ritchie MD, Crawford DC, Crane PK, Newton K, Li R, Mirel DB, Crenshaw A, Larson EB, Carlson CS, Jarvik GP; Electronic Medical Records and Genomics (eMERGE) Network.** Genetic variants associated with the white blood cell count in 13,923 subjects in the eMERGE Network. *Hum Genet* 131: 639–652, 2012. doi:10.1007/s00439-011-1103-9.
7. **Dillon SR, Sprecher C, Hammond A, Bilsborough J, Rosenfeld-Franklin M, Presnell SR, Haugen HS, Maurer M, Harder B, Johnston J, Bort S, Mudri S, Kuijper JL, Bukowski T, Shea P, Dong DL, Dasovich M, Grant FJ, Lockwood L, Levin SD, LeCiel C, Waggie K, Day H, Topouzis S, Kramer J, Kuestner R, Chen Z, Foster D, Parrish-Novak J, Gross JA.** Interleukin 31, a cytokine produced by activated T cells, induces dermatitis in mice. *Nat Immunol* 5: 752–760, 2004. [Erratum in *Nat Immunol* 6:114, 2005] 10.1038/ni1084.
8. **Dong S, Chen J.** SUMOylation of sPRDM16 promotes the progression of acute myeloid leukemia. *BMC Cancer* 15: 893, 2015. doi:10.1186/s12885-015-1844-2.
9. **Eicher JD, Chami N, Kacprowski T, Nomura A, Chen MH, Yanek LR, Tajuddin SM, Schick UM, Slater AJ, Pankratz N, Polfus L, Schurmann C, Giri A, Brody JA, Lange LA, Manichaikul A, Hill WD, Pazoki R, Elliot P, Evangelou E, Tzoulaki I, Gao H, Vergnaud AC, Mathias RA, Becker DM, Becker LC, Burt A, Crosslin DR, Lyytikäinen LP, Nikus K, Hernesniemi J, Kähönen M, Raitoharju E, Mononen N, Raitakari OT, Lehtimäki T, Cushman M, Zakai NA, Nickerson DA, Raffield LM, Quarells R, Willer CJ, Peloso GM, Abecasis GR, Liu DJ, Deloukas P, Samani NJ, Schunkert H, Erdmann J, Fornage M, Richard M, Tardif JC, Rioux JD, Dube MP, de Denus S, Lu Y, Bottinger EP, Loos RJ, Smith AV, Harris TB, Launer LJ, Gudnason V, Velez Edwards DR, Torstenson ES, Liu Y, Tracy RP, Rotter JI, Rich SS, Highland HM, Boerwinkle E, Li J, Lange E, Wilson JG, Mihailov E, Mägi R, Hirschhorn J, Metspalu A, Esko T, Vacchi-Suzzi C, Nalls MA, Zonderman AB, Evans MK, Engström G, Orho-Melander M, Melander O, O'Donoghue ML, Waterworth DM, Wallentin L, White HD, Floyd JS, Bartz TM, Rice KM, Psaty BM, Starr JM, Liewald DC, Hayward C, Deary IJ, Greinacher A, Völker U, Thiele T, Völzke H, van Rooij FJ, Uitterlinden AG, Franco OH, Dehghan A, Edwards TL, Ganesh SK, Kathiresan S, Faraday N, Auer PL, Reiner AP, Lettre G, Johnson AD; Global Lipids Genetics Consortium; CARDIoGRAM Exome Consortium; Myocardial Infarction Genetics Consortium.** Platelet-related variants identified by exomechip meta-analysis in 157,293 individuals. *Am J Hum Genet* 99: 40–55, 2016. doi:10.1016/j.ajhg.2016.05.005.
10. **Fukuda S, Schmid-Schönbein GW.** Regulation of CD18 expression on neutrophils in response to fluid shear stress. *Proc Natl Acad Sci USA* 100: 13152–13157, 2003. doi:10.1073/pnas.2336130100.
11. **Grove ML, Yu B, Cochran BJ, Haritunians T, Bis JC, Taylor KD, Hansen M, Borecki IB, Cupples LA, Fornage M, Gudnason V, Harris TB, Kathiresan S, Kraaij R, Launer LJ, Levy D, Liu Y, Mosley T, Peloso GM, Psaty BM, Rich SS, Rivadeneira F, Siscovick DS, Smith AV, Uitterlinden A, van Duijn CM, Wilson JG, O'Donnell CJ, Rotter JI, Boerwinkle E.** Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. *PLoS One* 8: e68095, 2013. doi:10.1371/journal.pone.0068095.
12. **Gusev A, Palamara PF, Aponte G, Zhuang Z, Darvasi A, Gregersen P, Pe'er I.** The architecture of long-range haplotypes shared within and across populations. *Mol Biol Evol* 29: 473–486, 2012. doi:10.1093/molbev/msr133.
13. **Halekoh U, Højsgaard S, Yan J.** The R package geepack for generalized estimating equations. *J Stat Softw* 15: 1–11, 2006. doi:10.18637/jss.v015.i02.
14. **Hanley JA, Negassa A, Edwardes MD, Forrester JE.** Statistical analysis of correlated data using generalized estimating equations: an orientation. *Am J Epidemiol* 157: 364–375, 2003. doi:10.1093/aje/kwf215.
15. **Kamatani Y, Matsuda K, Okada Y, Kubo M, Hosono N, Daigo Y, Nakamura Y, Kamatani N.** Genome-wide association study of hemato-

logical and biochemical traits in a Japanese population. *Nat Genet* 42: 210–215, 2010. doi:10.1038/ng.531.

16. **Keller MF, Reiner AP, Okada Y, van Rooij FJA, Johnson AD, Chen MH, Smith AV, Morris AP, Tanaka T, Ferrucci L, Zonderman AB, Lettre G, Harris T, Garcia M, Bandinelli S, Qayyum R, Yanek LR, Becker DM, Becker LC, Kooperberg C, Keating B, Reis J, Tang H, Boerwinkle E, Kamatani Y, Matsuda K, Kamatani N, Nakamura Y, Kubo M, Liu S, Dehghan A, Felix JF, Hofman A, Uitterlinden AG, van Duijn CM, Franco OH, Longo DL, Singleton AB, Psaty BM, Evans MK, Cupples LA, Rotter JI, O'Donnell CJ, Takahashi A, Wilson JG, Ganesh SK, Nalls MA; CHARGE Hematology; CO-GENT; BioBank Japan Project (RIKEN) Working Groups.** Trans-ethnic meta-analysis of white blood cell phenotypes. *Hum Mol Genet* 23: 6944–6960, 2014. doi:10.1093/hmg/ddu401.
17. **Kisucka J, Chauhan AK, Zhao BQ, Patten IS, Yesilaltay A, Krieger M, Wagner DD.** Elevated levels of soluble P-selectin in mice alter blood-brain barrier function, exacerbate stroke, and promote atherosclerosis. *Blood* 113: 6015–6022, 2009. doi:10.1182/blood-2008-10-186650.
18. **Kubota K, Shirakura T, Orui T, Muratani M, Maki T, Tamura J, Morita T.** [Changes in the blood cell counts with aging] [in Japanese]. *Nippon Ronen Igakkai Zasshi* 28: 509–514, 1991. doi:10.3143/geriatrics.28.509.
19. **Leslie R, O'Donnell CJ, Johnson AD.** GRASP: analysis of genotype-phenotype results from 1390 genome-wide association studies and corresponding open access database. *Bioinformatics* 30: i185–i194, 2014. doi:10.1093/bioinformatics/btu273.
20. **Li J, Glessner JT, Zhang H, Hou C, Wei Z, Bradfield JP, Mentch FD, Guo Y, Kim C, Xia Q, Chiavacci RM, Thomas KA, Qiu H, Grant SFA, Furth SL, Hakonarson H, Sleiman PMA.** GWAS of blood cell traits identifies novel associated loci and epistatic interactions in Caucasian and African-American children. *Hum Mol Genet* 22: 1457–1464, 2013. doi:10.1093/hmg/dds534.
21. **Liang KY, Zeger SL.** Longitudinal data analysis using generalized linear models. *Biometrika* 73: 13–22, 1986. doi:10.1093/biomet/73.1.13.
22. **Lin A, Wang RT, Ahn S, Park CC, Smith DJ.** A genome-wide map of human genetic interactions inferred from radiation hybrid genotypes. *Genome Res* 20: 1122–1132, 2010. doi:10.1101/gr.104216.109.
23. **Lo KS, Wilson JG, Lange LA, Folsom AR, Galarneau G, Ganesh SK, Grant SFA, Keating BJ, McCarroll SA, Mohler ER III, O'Donnell CJ, Palmas W, Tang W, Tracy RP, Reiner AP, Lettre G.** Genetic association analysis highlights new loci that modulate hematological trait variation in Caucasians and African Americans. *Hum Genet* 129: 307–317, 2011. doi:10.1007/s00439-010-0925-1.
24. **MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, Junkins H, McMahon A, Milano A, Morales J, Pendlington ZM, Welter D, Burdett T, Hindorff L, Flicek P, Cunningham F, Parkinson H.** The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* 45, *D1*: D896–D901, 2017. doi:10.1093/nar/gkw1133.
25. **Machiela MJ, Chanock SJ.** LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 31: 3555–3557, 2015. doi:10.1093/bioinformatics/btv402.
26. **McEver RP, Beckstead JH, Moore KL, Marshall-Carlson L, Bainton DF.** GMP-140, a platelet α-granule membrane protein, is also synthesized by vascular endothelial cells and is localized in Weibel-Palade bodies. *J Clin Invest* 84: 92–99, 1989. doi:10.1172/JCI114175.
27. **Montojo J, Zuberi K, Rodriguez H, Bader GD, Morris Q.** GeneMANIA: Fast gene network construction and function prediction for Cytoscape. *F1000Res* 3: 153, 2014. doi:10.12688/f1000research.4572.1.
28. **Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, Morris Q, Bader GD.** GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* 26: 2927–2928, 2010. doi:10.1093/bioinformatics/btq562.
29. **Mousas A, Ntritsos G, Chen MH, Song C, Huffman JE, Tzoulaki I, Elliott P, Psaty BM, Auer PL, Johnson AD, Evangelou E, Lettre G, Reiner AP; Blood-Cell Consortium.** Rare coding variants pinpoint genes that control human hematological traits. *PLoS Genet* 13: e1006925, 2017. doi:10.1371/journal.pgen.1006925.
30. **Okada Y, Hirota T, Kamatani Y, Takahashi A, Ohmiya H, Kumasaka N, Higasa K, Yamaguchi-Kabata Y, Hosono N, Nalls MA, Chen MH, van Rooij FJA, Smith AV, Tanaka T, Couper DJ, Zakai NA, Ferrucci L, Longo DL, Hernandez DG, Witteman JCM, Harris TB, O'Donnell CJ, Ganesh SK, Matsuda K, Tsunoda T, Tanaka T, Kubo M, Nakamura Y, Tamari M, Yamamoto K, Kamatani N.** Identification of nine novel loci associated with white blood cell subtypes in a Japanese population. *PLoS Genet* 7: e1002067, 2011. doi:10.1371/journal.pgen.1002067.
31. **Okada Y, Kamatani Y.** Common genetic factors for hematological traits in humans. *J Hum Genet* 57: 161–169, 2012. doi:10.1038/jhg.2012.2.
32. **van Oosterhout C.** A new theory of MHC evolution: beyond selection on the immune genes. *Proc Biol Sci* 276: 657–665, 2009. doi:10.1098/rspb.2008.1299.
33. **Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D.** Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909, 2006. doi:10.1038/ng1847.
34. **R Core Team.** R: A language and environment for statistical computing. [Online]. R Foundation for Statistical Computing, 2016. https://www.r-project.org/.
35. **van Rooij FJA, Qayyum R, Smith AV, Zhou Y, Trompet S, Tanaka T, Keller MF, Chang L-C, Schmidt H, Yang M-L, Chen M-H, Hayes J, Johnson AD, Yanek LR, Mueller C, Lange L, Floyd JS, Ghanbari M, Zonderman AB, Jukema JW, Hofman A, van Duijn CM, Desch KC, Saba Y, Ozel AB, Snively BM, Wu J-Y, Schmidt R, Fornage M, Klein RJ, Fox CS, Matsuda K, Kamatani N, Wild PS, Stott DJ, Ford I, Slagboom PE, Yang J, Chu AY, Lambert AJ, Uitterlinden AG, Franco OH, Hofer E, Ginsburg D, Hu B, Keating B, Schick UM, Brody JA, Li JZ, Chen Z, Zeller T, Guralnik JM, Chasman DI, Peters LL, Kubo M, Becker DM, Li J, Eiriksdottir G, Rotter JI, Levy D, Grossmann V, Patel KV, Chen CH, Ridker PM, Tang H, Launer LJ, Rice KM, Li-Gao R, Ferrucci L, Evans MK, Choudhuri A, Trompouki E, Abraham BJ, Yang S, Takahashi A, Kamatani Y, Kooperberg C, Harris TB, Jee SH, Coresh J, Tsai FJ, Longo DL, Chen YT, Felix JF, Yang Q, Psaty BM, Boerwinkle E, Becker LC, Mook-Kanamori DO, Wilson JG, Gudnason V, O'Donnell CJ, Dehghan A, Cupples LA, Nalls MA, Morris AP, Okada Y, Reiner AP, Zon LI, Ganesh SK; BioBank Japan Project.** Genome-wide trans-ethnic meta-analysis identifies seven genetic loci influencing erythrocyte traits and a role for *RBPMS* in erythropoiesis. *Am J Hum Genet* 100: 51–63, 2017. doi:10.1016/j.ajhg.2016.11.016.
36. **RStudio Team.** RStudio: Integrated development environment for R. [Online]. RStudio, Inc. RStudio, Inc., 2015. https://www.rstudio.com/.
37. **Schick UM, Jain D, Hodonsky CJ, Morrison JV, Davis JP, Brown L, Sofer T, Conomos MP, Schurmann C, McHugh CP, Nelson SC, Vadlamudi S, Stilp A, Plantinga A, Baier L, Bien SA, Gogarten SM, Laurie CA, Taylor KD, Liu Y, Auer PL, Franceschini N, Szpiro A, Rice K, Kerr KF, Rotter JI, Hanson RL, Papanicolaou G, Rich SS, Loos RJF, Browning BL, Browning SR, Weir BS, Laurie CC, Mohlke KL, North KE, Thornton TA, Reiner AP.** Genome-wide association study of platelet count identifies ancestry-specific loci in Hispanic/Latino Americans. *Am J Hum Genet* 98: 229–242, 2016. doi:10.1016/j.ajhg.2015.12.003.
38. **Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T.** Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498–2504, 2003. doi:10.1101/gr.1239303.
39. **Shin HY, Simon SI, Schmid-Schönbein GW.** Fluid shear-induced activation and cleavage of CD18 during pseudopod retraction by human neutrophils. *J Cell Physiol* 214: 528–536, 2008. doi:10.1002/jcp.21235.
40. **Sitlani CM, Rice KM, Lumley T, McKnight B, Cupples LA, Avery CL, Noordam R, Stricker BHC, Whitsel EA, Psaty BM.** Generalized estimating equations for genome-wide association studies using longitudinal phenotype data. *Stat Med* 34: 118–130, 2015. doi:10.1002/sim.6323.
41. **Tajuddin SM, Schick UM, Eicher JD, Chami N, Giri A, Brody JA, Hill WD, Kacprowski T, Li J, Lyytikäinen LP, Manichaikul A, Mihailov E, O'Donoghue ML, Pankratz N, Pazoki R, Polfus LM, Smith AV, Schurmann C, Vacchi-Suzzi C, Waterworth DM, Evangelou E, Yanek LR, Burt A, Chen MH, van Rooij FJA, Floyd JS, Greinacher A, Harris TB, Highland HM, Lange LA, Liu Y, Mägi R, Nalls MA, Mathias RA, Nickerson DA, Nikus K, Starr JM, Tardif JC, Tzoulaki I, Velez Edwards DR, Wallentin L, Bartz TM, Becker LC, Denny JC, Raffield LM, Rioux JD, Friedrich N, Fornage M, Gao H, Hirschhorn JN, Liewald DCM, Rich SS, Uitterlinden A, Bastarache L, Becker DM, Boerwinkle E, de Denus S, Bottinger EP, Hayward C, Hofman A, Homuth G, Lange E, Launer LJ, Lehtimäki T, Lu Y, Metspalu A, O'Donnell CJ, Quarells RC, Richard M, Torstenson ES, Taylor KD, Vergnaud AC, Zonderman AB, Crosslin DR, Deary IJ,**

Dörr M, Elliott P, Evans MK, Gudnason V, Kähönen M, Psaty BM, Rotter JI, Slater AJ, Dehghan A, White HD, Ganesh SK, Loos RJ, Esko T, Faraday N, Wilson JG, Cushman M, Johnson AD, Edwards TL, Zakai NA, Lettre G, Reiner AP, Auer PL. Large-scale exome-wide association analysis identifies loci for white blood cell traits and pleiotropy with immune-mediated diseases. *Am J Hum Genet* 99: 22–39, 2016. doi:10.1016/j.ajhg.2016.05.003.

42. **Tanaka M, Hirabayashi Y, Sekiguchi T, Inoue T, Katsuki M, Miyajima A.** Targeted disruption of oncostatin M receptor results in altered hematopoiesis. *Blood* 102: 3154–3162, 2003. doi:10.1182/blood-2003-02-0367.
43. **1000 Genomes Project Consortium; Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA.** A map of human genome variation from population-scale sequencing. *Nature* 467: 1061–1073, 2010. doi:10.1038/nature09534.
44. **Voorman A, Rice K, Lumley T.** Fast computation for genome-wide association studies using boosted one-step statistics. *Bioinformatics* 28: 1818–1822, 2012. doi:10.1093/bioinformatics/bts291.
45. **Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A, Mostafavi S, Montojo J, Shao Q, Wright G, Bader GD, Morris Q.** The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 38, *suppl_2*: W214–W220, 2010. doi:10.1093/nar/gkq537.
46. **Yamada Y, Matsui K, Takeuchi I, Oguri M, Fujimaki T.** Association of genetic variants with hypertension in a longitudinal population-based genetic epidemiological study. *Int J Mol Med* 35: 1189–1198, 2015. doi:10.3892/ijmm.2015.2151.
47. **Zhang J, Huang F-F, Wu D-S, Li W-J, Zhan H-E, Peng M-Y, Fang P, Cao P-F, Zhang M-M, Zeng H, Chen F-P.** SUMOylation of insulin-like growth factor 1 receptor, promotes proliferation in acute myeloid leukemia. *Cancer Lett* 357: 297–306, 2015. doi:10.1016/j.canlet.2014.11.052.
48. **Zhang Q, Putheti P, Zhou Q, Liu Q, Gao W.** Structures and biological functions of IL-31 and IL-31 receptors. *Cytokine Growth Factor Rev* 19: 347–356, 2008. doi:10.1016/j.cytogfr.2008.08.003.