Quantile-Specific Heritability of Mean Platelet Volume, Leukocyte Count, and Other Blood Cell Phenotypes

Paul T. Williams

Molecular Biophysics & Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Abstract

Introduction: “Quantile-dependent expressivity” occurs when the effect size of a genetic variant depends upon whether the phenotype (e.g., mean platelet volume, MPV) is high or low relative to its distribution. Methods: Offspring-parent regression slopes ($\beta_{\text{OP}}$) were estimated by quantile regression, from which quantile-specific heritabilities ($h^2$) were calculated ($h^2 = 2\beta_{\text{OP}}/[1 + r_{\text{spouse}}]$) for blood cell phenotypes in 3,929 parent-offspring pairs from the Framingham Heart Study. Results: Quantile-specific $h^2$ (±SE) increased with increasing percentiles of the offspring’s age- and sex-adjusted MPV distribution ($p_{\text{linear}} = 0.0001$): 0.48 ± 0.09 at the 10th, 0.53 ± 0.04 at the 25th, 0.70 ± 0.06 at the 50th, 0.74 ± 0.06 at the 75th, and 0.90 ± 0.12 at the 90th percentile. Quantile-specific $h^2$ also increased with increasing percentiles of the offspring’s white blood cell (WBC, $p_{\text{linear}} = 0.002$), monocyte ($p_{\text{linear}} = 0.01$), and eosinophil distributions ($p_{\text{linear}} = 0.0005$). In contrast, heritabilities of red blood cell (RBC) count, hematocrit (HCT), and hemoglobin (HGB) showed little evidence of quantile dependence. Quantile-dependent expressivity is consistent with gene-environment interactions reported by others, including (1) greater increases in WBC and PLT concentrations in subjects who are glutathione-S-transferase Mu1 ($GSTM1$) null homozygotes than $GSTM1$ sufficient when exposed to endotoxin; (2) significantly higher WBC count in AA homozygotes than carriers of the G-allele of the glutathione S-transferase P1 ($GSTP1$) rs1695 polymorphism at low but not high benzene exposure in shoe factory workers; (3) higher WBC counts in TT homozygotes than C-allele carriers of the interleukin-1β ($IL1B$) c.315C>T polymorphism after undergoing surgery for infective endocarditis but not before surgery. Discussion/Conclusion: Quantile-dependent expressivity may explain several purported gene-environment interactions involving blood cell phenotypes.

Keywords
Genetics · Leukocytes · Monocytes · Eosinophil · Platelet volume · Gene-environment interaction

Introduction

Hematological traits are medically important biomarkers that show moderate to strong heritability [1]. The commonest blood cells are red blood cells (RBCs, aka erythrocytes) whose primary role is to transport oxygen and carbon dioxide. RBCs represent about one half of the total blood volume [1]. Second most common are platelets (PLTs, aka thrombocytes) that vary by size and func-
tionality, whose primary function is clot formation to prevent bleeding [1]. Mean platelet volume (MPV) is a property determined when the PLT is produced by megakaryocytes [2]. A large MPV is associated with increased PLT aggregation, beta-thromboglobulin release, thromboxane synthesis, and greater risk for acute myocardial infarction, ischemic stroke, venous thrombosis, and re-stenosis following coronary angioplasty [3]. PLT count is a balance between thrombopoiesis and PLT senescence and consumption [4]. Whereas MPV reflects PLT size and functions, PLT count assesses PLT function, production, and aging [5]. It has been suggested that the inverse relationship between MPV and PLT count is due to the regulation of total PLT mass [2].

White blood cells (WBCs, aka leukocytes) play a fundamental role in the body’s innate and adaptive immune response to foreign pathogens [1]. They are derived from myeloid (neutrophils, basophils, eosinophils, and monocytes) and lymphoid lineages (lymphocytes). Total WBC count is a clinical indicator of inflammation status and predicts increased risk for total, cardiovascular, and cancer mortality. A $10^9/L$ increment in WBC count has been associated with a 32% increase in coronary heart disease risk and 20% increase in all-cause mortality [6, 7], even within the normal WBC count range. WBC counts increase with age, blood pressure, cigarette use, adiposity, and diabetes risk and in association with other plasma inflammatory markers [8]. Neutrophils represent between 50 and 70% of total WBCs and play an essential role in the innate immune system [1]. Lymphocytes represent 20–30% of WBCs and are part of the adaptive immune system. The remaining WBCs are monocytes (5–10%), eosinophils (1–5%), and basophils (0.5–1%) [1]. Additional blood cell phenotypes of clinical importance are hemoglobin (HGB) concentration, hematocrit (HCT, the fraction of the blood containing HGB), mean corpuscular hemoglobin (MCH, the average HGB mass per RBC), mean corpuscular hemoglobin concentration (MCHC, MCH concentration), mean corpuscular volume (MCV, average RBC volume), and RBC distribution width (RDW, a measure of the variation of RBC volume) [1].

“Quantile-dependent expressivity” refers to the phenotypic expression of a genetic variant being dependent on whether the trait is high or low relative to its distribution [9]. Prior analyses have shown that the heritability of adiposity [9, 10], coffee [11] and alcohol consumption [12], postprandial lipemia [13] and serum concentrations of total and HDL cholesterol [14–16], triglycerides [17], leptin [18], adiponectin [19], glucose and insulin [20], and C-reactive protein [21] all increase with increasing quantiles of each trait, whereas height and intakes of other macronutrients do not [9, 10, 12]. An important consequence of quantile-dependent expressivity is that sampling by characteristics that distinguish high-versus low-valued phenotypes is expected to produce different genetic effects [16]. These differences are traditionally ascribed to gene-environment, gene-drug, and gene-diet interactions when a simpler explanation of quantile-dependent expressivity might suffice. Included among these are 56 published purported interactions involving adiposity [10], 64 involving postprandial lipemia [13], 76 involving serum triglycerides [17], 21 involving total cholesterol [14], 88 involving HDL-cholesterol concentrations [15, 16], 45 involving fasting glucose [20], and 50 examples involving C-reactive protein concentrations [21].

This report applies quantile regression [22, 23] to offspring-parent and full-sib pairs from the Offspring and Third Generation Cohorts of the Framingham Heart Study [24, 25] to test whether the heritabilities of blood cell traits are constant throughout their range as traditionally assumed, or change systematically with increasing phenotype values, i.e., quantile-dependent expressivity [9]. Examples are presented of gene-environment interactions involving blood cell phenotypes that might be more simply explained by quantile-dependent expressivity.

Methods

The Framingham Study data were obtained from the National Institutes of Health FRAMCOHORT, GEN3, FRAMOFFSPRING Research Materials obtained from the National Heart, Lung, and Blood (NHLBI) Biologic Specimen and Data Repository Information Coordinating Center. The hypothesis tested not considered as part of the initial Framingham Study design and is exploratory. The Offspring Cohort included 5,124 adult children of the Original Cohort and their spouses who were initially examined between 1971 and 1975, reexamined 8 years later, and then every 3 to 4 years thereafter [24]. The Third Generation Cohort is composed of the children of the Offspring Cohort [25]. Subjects used in the current analyses were at least 16 years of age and were self-identified as non-Hispanic white.

Phlebotomy was performed on fasting participants who had rested for 5–10 min in a supine position, typically between 8:00 and 9:00 a.m. [26]. Specimens were stored at −80°C without freeze-thaw cycles until assay. The Wintrobe method was used to measure HCT [27]. Collected blood was spun at 5,000 rpm for 20 min in a balanced oxalate tube [26]. A calibrated scale was used to visually estimate the proportion of total blood volume due to RBCs [26]. MCV and MCH were the ratios of HCT and HBG to RBC count [26]. MPV, PLT, and neutrophil, lymphocyte, eosinophil, and monocyte counts were measured for examination 9 of the Offspring Cohort and examination 2 of the Third Generation Cohort. WBC counts were made for examination 9 of the Offspring Cohort and examination 2 of the Third Generation Cohort. WBC counts were made for examination 9 of the Offspring Cohort and examination 2 of the Third Generation Cohort.
narrow sense \(h^2\) was calculated as \(h^2 = \frac{\sigma^2_{sp}}{1 + r_{spouse}}\) and \(h^2 = 1 + 8\frac{\sigma^2_{spouse}}{1 + r_{spouse}}\) [29], where \(r_{spouse}\) is the spouse correlation.

Quantile regression uses all of the data to estimate the effect of the independent variable on each quantile of the dependent variable. Simultaneous quantile regression was performed using the sqreg command of Stata (version 11; StataCorp, College Station, TX, USA). One thousand bootstrap samples were drawn to estimate the variance-covariance matrix for the 91 quantile regression coefficients between the 5th and 95th percentiles of the offspring’s distribution [23]. Quantile-specific expressivity was assessed by (1) estimating quantile-specific \(\beta\) coefficients (± standard error [SE]) for the 5th, 6th,..., 95th percentiles of the sample distribution; (2) plotting the quantile-specific \(\beta\) coefficient versus the quantile of the trait distribution; and (3) testing whether the resulting graph was constant, or changed as linear, quadratic, or cubic functions of the percentile of the trait distribution using orthogonal polynomials [30].

Simulations of the test for trends under the traditional model (constant heritability for all percentiles) and quantile-dependent expressivity (increasing heritability with increasing percentiles of the trait distribution) have been provided previously [31]. Specifically, polygenic inheritance with constant heritability throughout the phenotype distribution (the traditional model) was simulated by creating a bivariate normal distribution from a linear combination of three independent normally distributed random variables. For heritability between zero and one, quantile regression analysis of 100,000 simulated data points yields parallel regression lines for all percentiles (quantiles) of the offspring values, i.e., no significant linear trend in the regression coefficients with increasing percentiles of the dependent variable (offspring phenotype). Other simulations showed quantile regression detected significant increases in \(h^2\) with increasing quantiles of the offspring distribution for data generated under quantile-dependent heritability.

Statistical analyses have been described in multiple reports [9–21] and summarized briefly here. Individual subject values are the average of the age- and sex-adjusted phenotypes over all available exams. Age and sex adjustment was performed separately in the Offspring and Third Generation Cohorts using standard least-squares regression with the following independent variables: father, mother, male × age, female × age, and female × age\(^2\) (constant heritability for all quantiles) and quantile-dependent expressivity. The results from several other published studies were reinterpreted from the perspective of quantile-dependent expressivity using genotype-specific means presented in the original articles or by extracting these values from published graphs [32, 33] using the Microsoft PowerPoint formatting palette as previously described [13]. Our interpretations of other studies are not necessarily those of the original authors.
Results

Table 1, which displays the sample characteristics, shows men tended to have higher RBC, HGB, HCT, and lower PLT. Estimated age- and sex-adjusted spouse correlations ($r_{\text{spouse}}$) were $-0.0063$ for MPV, $-0.0559$ for PLT, $0.1893$ for WBC $0.0290$ for neutrophil, $0.0215$ for lymphocyte, $0.0809$ for monocyte, and $0.0117$ for eosi-
phil count), 0.0259 for RBC, 0.1365 for MCV, 0.0678 for RDW, 0.0643 for HGB, 0.0976 for HCT, 0.0980 for MCH, and 0.0605 for MCHC. Table 2 presents the traditional point estimates of heritability that assume $h^2$ is constant throughout the phenotype distribution. All traits showed significant heritability, with MPV heritability being the strongest, and PLT, WBC, RBC, MCV, HCT, HGB, and MCH $h^2$ moderately strong. Variation in basophil counts was too limited for meaningful analysis.

**Platelets**

Figure 1a presents the offspring-parent regression lines at the 10th, 25th, 50th, 75th, and 90th percentiles of the offspring’s age- and sex-adjusted MPV distribution with their calculated heritabilities. The slopes ($\beta_{OP}$) get progressively steeper with increasing percentiles of the distribution, such that the slope at the 90th percentile was 86% greater than the slope at the 10th percentile ($p = 0.004$). These slopes along with those of the other percentiles between the 5th and 95th percentiles are presented in the quantile-specific heritability plot in Figure 1b. They show that MPV heritability increased linearly with increasing percentiles of the offspring’s MPV distribution (i.e., slope ± SE: 0.0024 ± 0.0006, $p_{\text{linear}} = 0.0001$ with no evidence of nonlinearity ($p_{\text{quadratic}} = 0.35, p_{\text{cubic}} = 0.48$).

Figure 1c shows the quantile-specific regression slopes for full sibs ($\beta_{FS}$) also increased linearly with increasing percentiles of the MPV distribution (i.e., slope ± SE: 0.0024 ± 0.0007, $p_{\text{linear}} = 0.0005, p_{\text{quadratic}} = 0.74, p_{\text{cubic}} = 0.35$). The slope at the 90th percentile was nearly twice as great (96% greater) as the slope at the 10th percentile ($p_{\text{difference}} = 1.6 \times 10^{-5}$). The quantile-specific $h^2$ estimated from $\beta_{OP}$ (represented by the dashed line) falls fully within the 95% confidence interval for the quantile-specific $h^2$ estimated from $\beta_{FS}$. Quantile-specific $h^2$ for each of the individual percentiles between the 5th and 95th percentiles were individually significant at $p \leq 1.6 \times 10^{-7}$ when estimated from $\beta_{OP}$ and $p \leq 6.2 \times 10^{-5}$ when estimated from $\beta_{FS}$. Figure 2a and Table 3 shows that the heritability for PLT also increased linearly with increasing percentiles of the PLT distribution when estimated from $\beta_{OP}$ (0.0025 ± 0.0007 increase per percent increment in the
PLT distribution, $p_{\text{linear}} = 0.0007$) and when estimated from $\beta_{\text{FS}}$ but not significantly so ($0.0015 \pm 0.0008$, $p_{\text{linear}} = 0.08$).

**Leukocytes**

Figure 2b displays the linear increase in WBC heritability as the offspring WBC increased from the 5th through the 95th percentiles ($0.0016 \pm 0.0005$, $p_{\text{linear}} = 0.02$) when calculated from $\beta_{\text{OP}}$. Table 3 shows the linear increase in WBC $h^2$ was further verified by heritability estimated from $\beta_{\text{FS}}$ ($0.0020 \pm 0.0005$, $p_{\text{linear}} = 0.0002$). Compared to the heritability at the 10th percentile, heritability at the 90th percentile of the WBC distribution was 2.2-fold greater when calculated from $\beta_{\text{OP}}$ ($p_{\text{difference}} = 0.003$) and 2.2-fold greater when calculated from $\beta_{\text{FS}}$ ($p_{\text{difference}} = 0.0001$). Moreover, Table 3 shows that $\beta_{\text{OP}}$-derived and $\beta_{\text{FS}}$-derived $h^2$ both show increasing heritability with increasing percentiles of the monocyte and eosinophil distribution. Quantile-specific heritability of neutrophil and lymphocyte increased linearly when calculated from siblings, albeit the increase in $\beta_{\text{OP}}$ for these leukocytes did not attain statistical significance.

**Other Hematological Parameters**

There was little evidence for quantile-dependent expressivity of RBC, MCV, RBC, HCT, HGB, or MCHC. MCH heritability may decrease with increasing percentiles of the MCH distribution.

**Log-Transformed Blood Cell Phenotypes**

The heritability of log transformed MPV increased with increasing concentrations of its distribution when estimated from $\beta_{\text{OP}}$ ($p_{\text{linear}} = 0.02$) and $\beta_{\text{FS}}$ ($p_{\text{linear}} = 0.06$). Similarly, the heritability of log-transformed eosinophil counts also increased with increasing concentrations of its distribution when estimated from $\beta_{\text{OP}}$ ($p_{\text{linear}} = 0.002$) and $\beta_{\text{FS}}$ ($p_{\text{linear}} = 9.2 \times 10^{-5}$). However, the transformation eliminated the significance increase in $\beta_{\text{OP}}$-derived and $\beta_{\text{FS}}$-derived heritability with increasing concentrations of PLT ($p_{\text{linear}} = 0.33$ and 0.35, respectively), WBC ($p_{\text{linear}} = 0.94$ and 0.31, respectively), neutrophil ($p_{\text{linear}} = 0.20$ and 0.72, respectively), lymphocyte ($p_{\text{linear}} = 0.59$ and 0.92, respectively), and monocyte counts ($p_{\text{linear}} = 0.77$ and 0.38, respectively).

**Gene-Environment Interactions**

Quantile-dependent expressivity may explain gene-environment interactions reported by others.

**Example 1**

The glutathione-S-transferase Mu1 (GSTM1) gene is one of several antioxidant genes affecting the systemic inflammatory response to environmental agents [32]. Approximately 40–50% of the population is homozygous for the GSTM1 gene deletion (null) polymorphism [34], which places them at increased risk for cardiovascular disease events from ambient and indoor air particulate matter, childhood wheezing from tobacco smoke, and acute exacerbation of asthma from ambient air O₃ [34, 35]. Dillon et al. [32] examined the WBC response to endotoxin by GSTM1 genotypes to assess the gene’s affect on systemic inflammation. They reported that when exposed to endotoxin, there were greater increases in WBC concentrations in subjects who were GSTM1 null homozygotes than in GSTM1 sufficient volunteers (Fig. 3a, histogram showing mean changes in WBC con-
centrations by genotype), suggesting that the GSTM1 null genotype is a risk factor, systemic inflammatory responses to environmental endotoxin. They also reported that GSTM1 null homozygotes showed significant increases in PLT concentrations when exposed to endotoxin that was not observed in GSTM1 sufficient volunteers (Fig. 3b). Alternatively, the associated line graph suggests the genetic effect size was accentuated at the

Table 3. Quantile regression analyses of offspring-parent and full-sib blood cell phenotypes from the Framingham Heart Study

| Phenotype       | Increase in slope per 1% increase in the offspring’s distribution | Difference in slope between the 90th and 10th percentiles |
|-----------------|---------------------------------------------------------------|-----------------------------------------------------------|
|                 | Linear effect | Nonlinear effects | Difference ± SE | p value | Linear effect | Nonlinear effects | Difference ± SE | p value |
|                 | slope ± SE  | quadratic p | cubic p |                           | slope ± SE  | quadratic p | cubic p |                           |
| MPV             |             |               |             |                           |             |               |             |                           |
| Offspring parent| 0.0024±0.0006| 0.35           | 0.48        | 0.208±0.072               | 0.004       |
| Full sibling    | 0.0024±0.0007| 0.74           | 0.35        | 0.235±0.054               | 1.6×10⁻⁵    |
| PLT             |             |               |             |                           |             |               |             |                           |
| Offspring parent| 0.0025±0.0007| 0.34           | 0.14        | 0.211±0.084               | 0.01        |
| Full sibling    | 0.0015±0.0008| 0.67           | 0.34        | 0.065±0.062               | 0.30        |
| WBC             |             |               |             |                           |             |               |             |                           |
| Offspring parent| 0.0016±0.0005| 0.55           | 0.71        | 0.157±0.053               | 0.003       |
| Full sibling    | 0.0020±0.0005| 0.97           | 0.03        | 0.192±0.050               | 0.0001      |
| Neutrophil count|             |               |             |                           |             |               |             |                           |
| Offspring parent| 0.0011±0.0007| 0.12           | 0.64        | 0.069±0.055               | 0.21        |
| Full sibling    | 0.0024±0.0007| 0.58           | 0.93        | 0.208±0.039               | 1.3×10⁻⁷    |
| Lymphocyte count|             |               |             |                           |             |               |             |                           |
| Offspring parent| 0.0012±0.0008| 0.16           | 0.19        | 0.129±0.068               | 0.06        |
| Full sibling    | 0.0024±0.0008| 0.80           | 0.87        | 0.197±0.104               | 0.06        |
| Monocyte count  |             |               |             |                           |             |               |             |                           |
| Offspring parent| 0.0014±0.0006| 0.01           | 0.54        | 0.128±0.046               | 0.005       |
| Full sibling    | 0.0019±0.0007| 0.67           | 0.16        | 0.182±0.067               | 0.007       |
| Eosinophil count|             |               |             |                           |             |               |             |                           |
| Offspring parent| 0.0030±0.0009| 0.38           | 0.64        | 0.283±0.060               | 2.7×10⁻⁶    |
| Full sibling    | 0.0040±0.0008| 9.8×10⁻⁷       | 0.32        | 0.374±0.105               | 0.0004      |
| RBC             |             |               |             |                           |             |               |             |                           |
| Offspring parent| 0.0007±0.0005| 0.13           | 0.23        | 0.096±0.041               | 0.02        |
| Full sibling    | 0.0007±0.0005| 0.79           | 0.47        | 0.058±0.056               | 0.30        |
| MCV             |             |               |             |                           |             |               |             |                           |
| Offspring parent| −0.0004±0.0004| 0.32           | 0.23        | −0.018±0.039               | 0.65        |
| Full sibling    | −0.0004±0.0005| 0.48           | 0.60        | −0.021±0.055               | 0.70        |
| RBC distribution|             |               |             |                           |             |               |             |                           |
| Offspring parent| −0.0002±0.0005| 0.70           | 0.38        | −0.034±0.041               | 0.41        |
| Full sibling    | 0.0003±0.0006| 0.57           | 0.39        | 0.032±0.082               | 0.70        |
| HCT             |             |               |             |                           |             |               |             |                           |
| Offspring parent| −0.0004±0.0005| 0.48           | 0.58        | −0.048±0.051               | 0.35        |
| Full sibling    | 0.0005±0.0004| 0.83           | 0.96        | 0.020±0.042               | 0.64        |
| HGB             |             |               |             |                           |             |               |             |                           |
| Offspring parent| 0.0001±0.0003| 0.68           | 0.78        | 0.012±0.035               | 0.73        |
| Full siblings   | 0.0006±0.0005| 0.99           | 0.82        | 0.041±0.052               | 0.43        |
| MCH             |             |               |             |                           |             |               |             |                           |
| Offspring parent| −0.0013±0.0005| 0.66           | 0.23        | −0.088±0.044               | 0.04        |
| Full siblings   | −0.0008±0.0005| 0.84           | 0.25        | −0.084±0.037               | 0.02        |
| MCHC            |             |               |             |                           |             |               |             |                           |
| Offspring parent| −0.0001±0.0004| 0.92           | 0.89        | −0.040±0.052               | 0.44        |
| Full sibling    | −0.0001±0.0004| 0.92           | 0.81        | −0.019±0.038               | 0.63        |

See Table 2 for abbreviations, Table 2 for sample sizes.
higher average WBC and PLT counts produced by endotoxin exposure, consistent with quantile-dependent expressivity. Specifically, the difference in the mean WBC concentrations between genotypes was greater after exposure to endotoxin when the overall WBC concentration was high ($7.27 \times 10^9/L$) than pre-exposure when overall WBC concentration was lower ($6.03 \times 10^9/L$). Some caution is warranted in interpreting these results given the limited size of the sample (17 \textit{GSTM1} null, 18 \textit{GSTM1} sufficient volunteers).

\textbf{Example 2}

Even at relatively low concentrations, benzene exposure is linked to reduced WBC counts [36], possibly due to the hematotoxicity of benzene metabolites rather than benzene per se [37]. Ye et al. [38] reported on WBC concentrations from worker routine medical evaluations at six shoe factories in Zhejiang Province, China. Benzene inhalation commonly occurs in these workers when glue brushes are dipped manually into open containers. Exposures as low as $<1$ p.p.m. of benzene in air are reported to significantly decrease WBC and PLT concentrations [39]. In Ye et al.’s data, average WBC counts in workers were greater at the lowest quartile of 16-month estimated cumulative benzene exposure than at higher exposure levels ($6.17 \pm 0.16$ vs. $5.42 \pm 0.09 \times 10^{-9}$).

The gene \textit{GSTP1} encodes the important antioxidant enzyme glutathione S-transferase Pi 1. Its rs1695 A>G polymorphism (aka rs947894, Ile105Val) is an A to G transition that decreases the protein’s substrate specificity and catalytic activity [40]. The rs1695 variant is the only common \textit{GSTP1} nonsynonymous polymorphism, with minor allele frequencies ranging from 17–44% [40]. The rs1695 A>G polymorphism lowers the benzene detoxification activity of glutathione S-transferase Pi 1.

Ye et al. compared mean WBC concentrations by rs1695 genotypes and quartiles of 16-month benzene exposure cross-sectionally [38]. The line graph of their data in Figure 4 shows WBC count was significantly greater in AA homozygotes than carriers of the G-allele of the \textit{GSTP1} rs1695 polymorphism at the lowest quartile of cumulative benzene exposure ($0.70 \pm 0.31 \times 10^{-9}$, $p = 0.02$) but not at higher cumulative benzene exposure ($-0.01 \pm 0.20 \times 10^{-9}$, $p_{\text{interaction}} = 0.05$), consistent with quantile-dependent expressivity. The greater efficacy of the rs1695 AA homozygotes vis-à-vis G-allele carriers appears to be eliminated at higher benzene exposure levels.

\textbf{Example 3}

Infective endocarditis is a microbiological infection of heart valve endothelial linings, primarily by \textit{Streptococcus}, \textit{Staphylococcus}, and \textit{Enterococcus} gram-positive bacteria, requiring surgical intervention in about one-half of
Quantile-Specific Blood Cell Heritability

Cases [41, 42]. Single nucleotide polymorphisms (SNP) are known to influence the course of sepsis and other bloodstream infections, but whether there are SNPs that modify cytokine host response to infective endocarditis is less studied.

Weinstock et al. [33] reported that TT homozygotes of the interleukin-1β (IL1B) c.315C>T polymorphism had higher WBC counts than either CT (p = 0.002) or CC (p = 0.02) genotypes after undergoing surgical intervention but not before. The histogram in Figure 5 suggests that this relates to a larger surgical effect on WBC count in TT than CT/CC genotypes, which could be the consequence of surgical stress. Similar results were observed for pre- and postoperative C-reactive protein levels. Alternatively, the associated line graph shows that the larger postoperative genetic effect size occurred in the context of a higher postoperative mean WBC count, and the smaller preoperative genetic effect size in the context of the smaller preoperative mean count. These results are important given that abnormal WBC counts have been shown to predict higher infective endocarditis mortality [43].

Discussion

MPV had the highest overall (i.e., traditional) heritability of the hematological traits analyzed (Table 2). Moreover, Figure 1 showed that quantile-specific MPV heritability increased substantially with increasing MPV level. Genetic effects as measured by heritability were 86–96% greater at the 90th percentiles of the MPV distribution than at the 10th percentile. Figure 2a and Table 3 present evidence for quantile-dependent expressivity of PLT count as well. When estimated by βOP, PLT h² was over 3-fold greater at the 90th than at the 10th percentile of its distribution. Figure 2a and Table 3 also suggest that heritabilities for WBC, monocyte, and eosinophil counts increase linearly with increasing percentiles of their distributions and provide evidence for accentuated heritability of neutrophil and lymphocyte counts at their higher concentrations. An important consequence of quantile-dependent expressivity is that selecting subjects from the high versus low portions of MPV, PLT, or WBC distributions is expected to yield different estimates of the genetic effect size. In that regard, quantile-dependent ex-
pressivity may provide an alternative explanation to purported environment interactions, as illustrated in three examples.

Utility of Reporting Genotype Differences in Controls
Quantile-dependent expressivity is a novel hypothesis. Therefore, it is unsurprising that many reports lack the information required to access its applicability, namely, genotype-specific mean WBC, leukocyte, and PLT counts by environmental health condition. For example, Yu et al. [44] reported that patients with virus-induced encephalitis showed significantly increasing concentrations in WBC, neutrophil, lymphocyte, and monocytes per A-allele of the interleukin 4 (IL-4) rs2243288 polymorphism and per T-allele of interleukin 10 (IL-10) rs1800872 polymorphism. Quantile-dependent expressivity would predict that the genetic effect size of these polymorphisms should be greater at the higher mean concentrations of the patients than healthy controls for WBC (patient vs. control mean concentration: 6.95 ± 0.08 vs. 4.55 ± 0.08 × 10⁹), neutrophil (9.38 ± 0.21 vs. 3.88 ± 0.19 × 10⁹), lymphocyte (5.49 ± 0.10 vs. 2.37 ± 0.06 × 10⁹), and monocyte counts (3.70 ± 0.03 vs. 0.58 ± 0.15 × 10⁹). However, this cannot be tested because the genotype differences in controls were not reported. Another study, by Michel et al. [45], reported that AG heterozygotes of the toll-like receptor 4 (TLR4) showed smaller increases in WBC count 6 h after inhaling lipopolysaccharide than AA homozygotes, but in the absence of baseline or 6-h average WBC concentrations it cannot be determined whether the genotype differences were accentuated at the higher mean concentrations at 6 h versus baseline. Similarly, quantile-dependent expressivity would be expected to accentuate genetic effects when average MPV is increased in association with hypertension, diabetes, obesity, hypercholesterolemia, smoking, and aspirin and diuretic use [46]. It also predicts smaller genetic effects for PLT concentrations for thrombocytopenia (<150,000 PLTs per microliter of circulating blood) and different genotype-specific PLT changes when corticosteroids are used to treat thrombocytopenia.

Heritability
Estimated heritability lacks specificity. Nevertheless, estimated heritability is important because that only 1.8% of the total WBC phenotype variance [47], and <10% of the estimated heritability of WBC traits, for example, are attributable to specific loci in populations of European ancestry [48]. Despite their simplicity, our estimates of heritability using offspring-parent and full-sib regression slopes and Falconer’s formula are not inconsistent with the highly variable estimates published by others using alternative estimation procedures. For example, published heritability estimates for WBC show 2-fold variation (0.35–0.71 [49–57]) with at least one study reporting significant dominance [50]. Moderately high heritabilities are also reported for lymphocyte (0.41 [54], 0.71 [57]), neutrophil (0.40 [54], 0.67 [57]), monocyte (0.57 [54], 0.66 [57]), and eosinophil counts (0.24 [58], 0.44 [54], 0.69 [57]). Strong heritability is reported for MPV (0.55–0.88 [52–55]), and PLT (0.53–0.86 [50, 52–57, 59]), and moderate to strong heritability for RBC (0.42–0.75 [50, 53–57]) HCT (0.32–0.65 [53–55]), MCV (0.20–0.96 [50, 53–57]), RDW (0.37 [53]), and MCH (0.39–0.91 [50, 53–56]) but not MCHC (0.00–0.46 [53–56]). Except for RBC, Evans et al.’s study of hematological traits in 392 twin pairs failed to identify effects due to common environment [57]. Thus, heritabilities of blood cell phenotypes are well studied; however, none of the prior studies considered quantile-specific effects, which could explain some of the variation between studies.

Data Transformation
Untransformed blood cell phenotypes were analyzed so that the findings apply to the traits as originally measured and because quantile regression and its bootstrap-derived standard errors do not require normally distributed data. Decisions to logarithmically transform WBC and other hematological phenotypes concentration have always been based on the requirements of the parametric statistical tests rather than a biological rationale. Genome-wide association studies are often performed on log WBC, MPV, PLT, and leukocyte phenotypes [47, 60]; however, this statistical accommodation may work against the goal of identifying SNP associations. Specifically, the logarithmic transformation emphasizes differences at lower phenotype values and de-emphasizes differences for higher-valued phenotypes, whereas our results suggest the largest genetic effects are at the highest MPV, PLT, and leukocyte concentrations.

Limitation
There are important limitations to our analyses. MPV measurements can be affected by pre-analytic factors including blood tube selection, processing, and equipment [61]. Our analyses do not include adjustment for covariates except for age and sex. Our analyses are limited to subjects of European ancestry and therefore exclude those of African ancestry for whom the rs2814778 variant
accounts for a substantial portion of their WBC heritability [62]. The analyses did not consider confounding variables that might also cluster within families and contribute to offspring-parent and sibling concordance. Finally, we note that the use of Falconer’s formula may underestimate the complexity of genetic and nongenetic contributions to blood cell phenotypes. In this regard, not allowing spouse concordance to vary by quantile is a potential weakness of the method. However, regression analysis showed no evidence that the offspring phenotypes affected the relationship between spouses for MPV ($p = 0.36$), PLT ($p = 0.31$), WBC ($p = 0.25$), monocyte ($p = 0.61$), or eosinophil distribution ($p = 0.68$). In addition, the larger $\beta_{FS}$-derived heritability than $\beta_{OP}$-derived heritability estimates of Table 2 may reflect dominance variance or additional family associations.

**Conclusion**

Our analyses suggest that the phenotypic effect of genetic variants may depend upon whether leukocyte concentrations, PLT concentrations, and MPV levels are high or low relative to their distribution. This could potentially contribute wholly or in part to gene-environment interactions involving blood cell phenotypes. Specifically, quantile-dependent expressivity may explain reported interactions between (1) greater increases in WBC and PLT concentrations in subjects who are glutathione-S-transferase M1 (GSTM1) null homozygotes than GSTM1 sufficient when exposed to endotoxin; (2) significantly higher WBC count in AA homozygotes than carriers of the G-allele of the glutathione S-transferase P1 (GSTP1) rs1695 polymorphism at low but not high benzene exposure in shoe factory workers; (3) higher WBC counts in TT homozygotes than C-allele carriers of the interleukin-1β (IL1B) c.315C>T polymorphism after undergoing surgery for infective endocarditis but not before surgery. Quantile-dependent expressivity has been previously shown for a number of important cardiovascular [9–21, 31] and pulmonary risk factors [63]. The current results extend the finding of quantile dependence to heritability in important hematological phenotypes.

**Acknowledgments**

We appreciate the contribution of the clinic, staff, and investigator who initially collected, organized, and documented the Framingham Heart Study family data.

**Statement of Ethics**

Lawrence Berkeley National Laboratory Human Subjects Committee (HSC) approved the analyses of these data for protocol “gene-environment interaction versus quantile-dependent penetrance of established SNPs (107H021).” LBNL holds Office of Human Research Protections Federal-Wide Assurance number FWA 00006253 (approval number: 107H021-13MR20). All surveys were conducted under the direction of the Framingham Heart Study human use committee guidelines, with written signed informed consent from all participants or parent and/or legal guardian if <18 years of age.

**Conflict of Interest Statement**

There are no conflicts of interest to report.

**Funding Sources**

This research was supported by NIH Grant R21ES020700 from the National Institute of Environmental Health Sciences. The funders had no involvement in the preparation of this manuscript.

**Author Contributions**

Dr. Paul T. Williams is solely responsible for the content of this paper (statistical analyses, authorship, and figure preparation).

**Data Availability Statement**

The data are not being published in accordance with the data use agreement between the NIH National Heart, Lung, and Blood Institute and Lawrence Berkeley National Laboratory. However, the data that support the findings of this study are available from NIH National Heart, Lung, and Blood Institute Biologic Specimen and Data Repository Information Coordinating Center directly through the Website https://biolincc.nhlbi.nih.gov/my/submitted/request/. Further inquiries can be directed to the corresponding author.

**References**

1. Okada Y, Kamatani Y. Common genetic factors for hematological traits in humans. J Hum Genet. 2012;57(3):161–9.
2. Korniluk A, Koper-Lenkiewicz OM, Kamińska J, Kemona H, Dymicka-Piekarska V. Mean Platelet Volume (MPV): new perspectives for an old marker in the course and prognosis of inflammatory conditions. Mediators Inflamm. 2019;2019:9213074.
3. Nugent D, Kunicki T. Platelet genomics: the role of platelet size and number in health and disease. Platelets. 2017;28(1):27–33.
4. Daly ME. Determinants of platelet count in humans. Haematologica. 2011;96(1):10–3.
5 Du J, Wang Q, He B, Liu P, Chen JY, Quan H, et al. Association of mean platelet volume and platelet count with the development and prognosis of ischemic and hemorrhagic stroke. Int J Lab Hematol. 2016;38(3):233–9.
6 de Labry LO, Campion EW, Glynn RJ, Vokonas PS. White blood cell count as a predictor of mortality: results over 18 years from the normative aging study. J Clin Epidemiol. 1990;43(2):153–7.
7 Kannel WB, Anderson K, Wilson PW. White blood cell count and cardiovascular disease. Insights from the Framingham study. JAMA. 1992;267(9):1253–6.
8 Gkrania-Klotsas E, Ye Z, Cooper AJ, Sharp SJ, Luben R, Biggs ML, et al. Differential white blood cell count and type 2 diabetes: systematic review and meta-analysis of cross-sectional and prospective studies. PLoS One. 2010;5(10):e13405.
9 Williams PT. Quantile-specific penetrance of genes affecting lipoproteins, adiposity and height. PLoS One. 2012;7(1):e28764.
10 Williams PT. Quantile-dependent heritability of computed tomography, dual-energy x-ray absorptiometry, anthropometric, and biochemical measures of adiposity. Int J Obesity. 2020;44(10):2101–12.
11 Williams PT. Quantile-specific heritability may account for gene-environment interactions involving coffee consumption. Behav Genet. 2020;50:119–26.
12 Williams PT. Quantile-specific heritability of intakes of alcohol but not other macronutrients. Behav Genet. 2020;50:332–45.
13 Williams PT. Quantile-dependent expressivity of postprandial lipemia. PLoS One. 2020;15:e0229495.
14 Williams PT. Quantile-specific heritability of total cholesterol and its pharmacogenetic and nutrigenetic implications. Int J Cardiol. 2021;327:185–92.
15 Williams PT. Quantile-specific heritability of high-density lipoproteins with implications for precision medicine. J Clin Lipidol. 2020;14:448–58.e60
16 Williams PT. Quantile-dependent expressivity and gene-lifestyle interactions involving high-density lipoprotein cholesterol. Lifestyle Genom. 2021;14:1–19.
17 Williams PT. Gene-environment interactions due to quantile-specific heritability of triglyceride and VLDL concentrations. Sci Rep. 2020;10:14868.
18 Williams PT. Quantile-specific heritability of sibbling leptin concentrations and its implications for gene-environment interactions. Sci Rep. 2020;10:22152.
19 Williams PT. Quantile-dependent expressivity of plasma adiponectin concentrations may explain its sex-specific heritability, gene-environment interactions, and genotype-specific response to postprandial lipemia. PeerJ. 2020;8:e10099.
20 Williams PT. Quantile-dependent heritability of glucose, insulin, proinsulin, insulin resistance, and glycated hemoglobin. Lifestyle Genom. 2022;5(1):10–34.
21 Williams PT. Quantile-dependent expressivity of serum C-reactive protein concentrations in family sets. PeerJ 2021;9:e10914.
22 Koenker R, Hallock KF. Quantile regression. J Econ Perspect. 2001;15(4):143–56.
23 Gould WW. Quantile regression with bootstrapped standard errors. Stata Tech Bull. 1992;9:19–21.
24 Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. Am J Epidemiol. 1979;110(3):281–90.
25 Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, et al. The third generation cohort of the National Heart, Lung, and Blood Institute’s Framingham Heart study: design, recruitment, and initial examination. Am J Epidemiol. 2007;165(11):1328–35.
26 Yang Q, Kathiresan S, Lin JP, Tofler GH, O’Donnell CJ. Genome-wide association and linkage analyses of hemostatic factors and hematological phenotypes in the Framingham Heart study. BMC Med Genet. 2007;8(Suppl 1):S12.
27 The principles and technic of blood examination. Philadelphia, PA: Lea & Febiger; 1961.
28 Karlin S, Cameron EC, Williams PT. Sibling and parent-offspring correlation estimation with variable family size. Proc Natl Acad Sci U S A. 1981;78(5):2664–8.
29 Falconer DS, Mackay TFC. Introduction to quantitative genetics. 4th ed. Harlow, UK: Longmans Green; 1996.
30 Winer BJ, Brown DR, Michels KM. Statistical principles in experimental design. 3rd ed. New York: McGraw-Hill; 1991.
31 Williams PT. Quantile-specific heritability of serum growth factor concentrations. Growth Factors. 2021;39(1–6):45–58.
32 Dillon MA, Harris B, Hernandez ML, Zou B, Reed WC, Lubenberg PA, et al. Enhancement of systemic and sputum granulocyte response to inhaled endotoxin in people with the GSTM1 null genotype. Occup Environ Med. 2011;68(10):783–5.
33 Weinstock M, Grimm I, Dreier J, Knabbe C, Grossmann V, et al. Differential white blood cell count and cardiovascular disease. JAMA. 1992;267(9):1253–6.
34 London SJ, Ruchirawat M, Biomarkers in population studies: environmental mutagenesis and risk for cancer. Rev Environ Health. 2009;24(2):117–27.
35 Ye LI, Zhang GH, Huang JW, Li Y, Zheng GQ, Zhang DT, et al. Are polymorphisms in metabolism protective or a risk for reduced white blood cell counts in a Chinese population with low occupational benzene exposure? Int J Occup Environ Health. 2015;21(3):232–40.
36 Lan Q, Zhang L, Li G, Vermeulen R, Weinberg RS, Dosemeci M, et al. Hematotoxicity in workers exposed to low levels of benzene. Science. 2004;306(5702):1774–6.
37 Agúndez JAG, García-Martín E, Martínez C, Benito-León J, Millán-Pascual J, Díaz-Sánchez M, et al. The GSTP1 gene variant rs1695 is not associated with an increased risk of multiple sclerosis. Cell Mol Immunol. 2015;12(6):777–9.
38 Michel O, LeVan TD, Stern D, Dentener M, Thorn J, Gnat D, et al. Systemic responsive- ness to lipopolysaccharide and polymyeloneuropathy in the toll-like receptor 4 gene in human beings. J Allergy Clin Immunol. 2003;112(5):923–9.
39 Panova-Noeva M, Schulz A, Hermanns MI, Grossmann V, Pefani E, Spronk HM, et al. Sex-specific differences in genetic and nongenetic determinants of mean platelet volume: results from the Gutenberg health study. Blood. 2016;127(2):251–9.
40 Panova-Noeva M, Scherpf UM, Eicher JD, Chami N, Giri A, Brody JA, et al. Large-scale exome-wide association analysis identifies loci for white blood cell traits and pleiotropy with immune-mediated diseases. Am J Hum Genet. 2016;99(1):22–39.
41 Nalls MA, Couper DJ, Tanaka T, van Rooij FJ, Chen MH, Smith AV, et al. Multiple loci are associated with white blood cell phenotypes. PLoS Genet. 2011;7(6):e1002113.
49 Pankow JS, Folsom AR, Cushman M, Borecki IB, Hopkins PN, Eckfeldt JH, et al. Familial and genetic determinants of systemic markers of inflammation: the NHLBI family heart study. Atherosclerosis. 2001;154(3):681–9.

50 Pilia G, Chen WM, Scuteri A, Orrù M, Albai G, Dei M, et al. Heritability of cardiovascular and personality traits in 6, 148 Sardinians. PLoS Genet. 2006;2(8):e132.

51 Cheng YC, Kao WH, Mitchell BD, Sharrett AR, Ryan KA, Vogel RA, et al. Genetic effects on postprandial variations of inflammatory markers in healthy individuals. Obesity. 2010; 18(7):1417–22.

52 Bray PF, Mathias RA, Faraday N, Yanek LR, Fallin MD, Herrera-Galeano JE, et al. Heritability of platelet function in families with premature coronary artery disease. J Thromb Haemost. 2007;5(8):1617–23.

53 Marroni F, Graziol D, Pattaro C, Devoto M, Pramstaller P. Estimates of genetic and environmental contribution to 43 quantitative traits support sharing of a homogeneous environment in an isolated population from South Tyrol, Italy. Hum Hered. 2008;65(3):175–82.

54 Traglia M, Sala C, Masciullo C, Cverhova V, Lori F, Pistis G, et al. Heritability and demographic analyses in the large isolated population of Val Borbera suggest advantages in mapping complex traits genes. PLoS One. 2009;4(10):e7554.

55 Whitfield JB, Martin NG. Genetic and environmental influences on the size and number of cells in the blood. Genet Epidemiol. 1985;2:133–44.

56 Garner C, Tatu T, Reittie JE, Littlewood T, Darley J, Cervino S, et al. Genetic influences on F cells and other hematologic variables: a twin heritability study. Blood. 2000;95(1):342–6.

57 Evans DM, Frazer IH, Martin NG. Genetic and environmental causes of variation in basal levels of blood cells. Twin Res. 1999;2(4):250–7.

58 Dal Colletto GM, Fulker DW, Barretto OC, Kolya M. Genetic and environmental effects on blood cells. Acta genet med gemellol (roma). Acta Genet Med Gemellol. 1993;42:425–52.

59 Biino G, Balduini CL, Casula L, Cavallo P, Vaccargiu S, Parracciani D, et al. Analysis of 12, 517 inhabitants of a Sardinian geographic isolate reveals that predispositions to thrombocytopenia and thrombocytosis are inherited traits. Haematologica. 2011;96(1):96–101.

60 Okada Y, Hirota T, Kamatani Y, Takahashi A, Ohmiya H, Kumasaka N, et al. Identification of nine novel loci associated with white blood cell subtypes in a Japanese population. PLoS Genet. 2011;7(6):e1002067.

61 Lance MD, Sloep M, Henskens YM, Marcus MA. Mean platelet volume as a diagnostic marker for cardiovascular disease: drawbacks of preanalytical conditions and measuring techniques. Clin Appl Thromb Hemost. 2012;18(6):561–8.

62 Reiner AP, Lettre G, Nalls MA, Ganesh SK, Mathias R, Austin MA, 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.

63 Williams PT. Spirometric traits show quantile-dependent heritability, which may contribute to their gene-environment interactions with smoking and pollution. PeerJ. 2020;8:e9145.