Published in final edited form as:

Thromb Haemost. 2015 April ; 113(4): 772–781. doi:10.1160/TH14-08-0670.

A genome-wide association study of heparin-induced thrombocytopenia using an electronic medical record

Jason H Karnes¹, Robert M Cronin¹, Jerome Rollin², Alexander Teumer³, Claire Pouplard², Christian M Shaffer¹, Carmelo Blanquicett¹, Erica A Bowton¹, James D Cowan¹, Jonathan D Mosley¹, Sara L Van Driest¹, Peter E Weeke^{1,4}, Quinn S Wells¹, Tamam Bakchoul³, Joshua C Denny¹, Andreas Greinacher³, Yves Gruel², and Dan M Roden¹

¹Vanderbilt University School of Medicine, Nashville, TN

²UMR CNRS 7292 & University Hospital of Tours, Tours Cedex, France

³University of Greifswald, Greifswald, Germany

⁴Copenhagen University Hospital, Gentofte, Denmark

Abstract

Heparin-induced thrombocytopenia (HIT) is an unpredictable, potentially catastrophic adverse effect of heparin treatment resulting from an immune response to platelet factor 4 (PF4)/heparin complexes. No genome-wide evaluations have been performed to identify potential genetic influences on HIT. Here, we performed a genome-wide association study (GWAS) and candidate gene study using HIT cases and controls identified using electronic medical records (EMRs) coupled to a DNA biobank and attempted to replicate GWAS associations in an independent cohort. We subsequently investigated influences of GWAS-associated single nucleotide polymorphisms (SNPs) on PF4/heparin antibodies in non-heparin treated individuals. In a recessive model, we observed significant SNP associations (OR 18.52 [6.33–54.23], $p=3.18 \times 10^{-9}$) with HIT near the T-Cell Death-Associated Gene 8 (*TDAG8*). These SNPs are in linkage disequilibrium with a missense *TDAG8* SNP. *TDAG8* SNPs trended toward an association with HIT in replication analysis (OR 5.71 [0.47–69.22], $p=0.17$), and the missense SNP was associated with PF4/heparin antibody levels and positive PF4/heparin antibodies in non-heparin treated patients (OR 3.09 [1.14–8.13], $p=0.02$). In the candidate gene study, SNPs at *HLA-DRA* were nominally associated with HIT (OR 0.25 [0.15–0.44], $p=2.06 \times 10^{-6}$). Further study of *TDAG8* and *HLA-DRA* SNPs is warranted to assess their influence on the risk of developing HIT.

Corresponding Author: Dan M Roden, M.D., MRB4 1285B, 2215B Garland Avenue, Nashville, TN, 37232-0575, Telephone: 615-322-0067, Fax: 615-343-4522, dan.roden@Vanderbilt.Edu.

Authorship Contributions

For the discovery GWAS cohort, JHK, RMC, EAB, JDM, SVD, PEW, QSW, JCD, and DMR designed the study, interpreted results; JHK, RMC, CB, EAB, and JCD designed case and controls algorithms, reviewed electronic medical records, and adjudicated case and control subjects; JHK, CMS, EAB, JDM, and JCD performed statistical analysis and provided bioinformatic support. For the SHIP study, AG designed the study; AT, AG, and TB interpreted results; AT performed statistical analysis. For the independent replication population, YG designed the study; JR, CP, and YG interpreted results; JR and JHK performed statistical analysis. All authors helped to write the manuscript.

Disclosure of Conflicts of Interest

The authors declare no competing financial interests.

Introduction

Heparin-induced thrombocytopenia (HIT) is an antibody-mediated condition of platelet activation in patients receiving unfractionated heparin (UFH) and low molecular weight heparin (LMWH)(1). HIT develops in up to 0.5–5% of patients treated with heparin anticoagulants, has a greater than 30% mortality rate, and results in catastrophic thromboembolic complications, including life- and limb-threatening thrombosis(2–5). Further complicating the use of heparin anticoagulants, prevention of HIT-related thrombosis is currently possible only after manifestations of HIT are evident and the disease process has already begun(6, 7). The inability to predict HIT thus represents a liability with heparin administration and *a priori* identification of patients with a high HIT risk could allow for prevention of HIT.

HIT is closely associated with the development of antibodies to complexes of heparin and platelet factor 4 (PF4), a protein normally found in the alpha granules of platelets(8). Although several studies have identified genetic polymorphisms such as the Fc receptor RIIA (*FCGR2A*) –H131R and Fc receptor RIIIA (*FCGR3A*) –F158V polymorphisms(9–13) associated with HIT, none have identified genetic risk factors that can reliably predict HIT. Other work has identified genomic predictors of diverse types of immune-mediated adverse drug reactions(14–16), suggesting that strong genetic influences on HIT risk might still be identified. This is of special interest, as HIT is an example for a distinct group of immune reactions, which fulfills neither the criteria of a primary nor the criteria of a typical secondary immune reaction(17–19). The clinical spectrum of HIT includes distinct stages of progression including 1) lack of PF4/heparin antibodies; 2) presence of PF4/heparin antibodies without thrombocytopenia; 3) PF4/heparin antibodies with thrombocytopenia; and 4) PF4/heparin antibodies with thrombocytopenia and thrombosis.(3, 20) As a result, single nucleotide polymorphisms (SNPs) may exert their effects only at distinct stages in the progression of HIT and different SNPs may be predictors of PF4/heparin antibody formation, thrombocytopenia, or thrombosis.

The objective of this study was to identify genetic predictors of HIT using genome-wide association study (GWAS) and candidate gene approaches. We conducted this study using HIT cases and heparin-exposed controls identified in an electronic medical record (EMR) linked to a DNA repository. Associated SNPs from the HIT GWAS were genotyped in an independent replication cohort of HIT cases and heparin-treated controls. We also tested whether GWAS-identified SNPs were associated with PF4/heparin antibody levels in non-heparin treated patients.

Materials and Methods

Study Populations

BioVU (Discovery Population)—The discovery population was identified in BioVU, the Vanderbilt DNA databank that links DNA extracted from discarded blood samples to de-identified EMRs(21). This study was approved by the Institutional Review Board at Vanderbilt University as described previously(21). We developed an algorithm to identify individuals with a diagnosis of HIT in the EMR. The algorithm used ICD9 diagnostic codes,

lab results, and natural language processing(22, 23) to analyze narrative text in order to identify patients treated with UFH or LMWH who tested positive for PF4/heparin antibodies (optical density [OD] level >0.7 absorbance units [AU]). UFH or LMWH-treated controls were also identified in the EMR using a control algorithm that employed similar techniques. HIT cases were confirmed by chart review by at least two physician or pharmacist reviewers. HIT cases were scored with the 4Ts scoring system, which estimates the likelihood of HIT based on timing, magnitude, and potential causes of thrombocytopenia and presence of thrombosis and with a maximum 4Ts score of eight(6, 24). Additional information regarding the BioVU resource and case and control algorithms is available in the Supplemental Materials.

Replication of GWAS SNPs—The patient group for replication has been previously described(11, 13) and additional details regarding the replication study population are detailed in the Supplemental Materials. In brief, the replication cohort was comprised of 94 Caucasian individuals with definite HIT who developed delayed-onset thrombocytopenia and/or a significant decrease in platelet count after a median duration of heparin treatment of 10 days. Both PF4-specific ELISA (HAT[®] GTI, Brookfield, WI) and serotonin release assay were positive in all patients. The first control group (Ab^{neg}) consisted of 178 patients who had undergone cardiopulmonary bypass surgery and received high doses of UFH but did not develop an immune response to heparin. The second control group (Ab^{pos}) consisted of 154 patients who had also undergone cardiopulmonary bypass surgery with heparin treatment. All had developed significant levels of PF4-specific antibodies but none had a positive serotonin release assay and none had significant abnormal evolution in platelet count in the post-operative period. Blood samples were collected after obtaining informed consent according to the Helsinki declaration principles. Both the University of Tours ethics committee and the Ministry of Research had previously approved collection of DNA from patients for genetic studies (Agreement number DC2008-308).

SNP Association with PF4/heparin Antibody Development—The Study of Health in Pomerania (SHIP) is a population-based project in West Pomerania, Germany that has been previously described(25–27). SHIP participants were recruited from population registries by written invitation regardless of previous anticoagulant treatment. Additional details regarding the SHIP population are available in the Supplemental Materials. The SHIP population comprised 4,308 participants who had sera collected and tested for anti-PF4/heparin IgG, IgA, and IgM antibodies by enzyme immunoassay(28), including a high heparin inhibition step as previously described(29). Reproducible optical density levels (ODs)>0.5 that were inhibited by high heparin (100 IU/mL) 40% were considered positive. All subjects gave their written informed consent in accordance with the Declaration of Helsinki. The study protocols were approved by the local ethics committee of Greifswald University.

Genotyping, Imputation, and Quality Control

BioVU (Discovery Population)—Samples were genotyped using the Illumina[®] HumanOmni1-QUAD and HumanOmni5-QUAD BeadChip genome-wide platforms. The discovery GWAS was performed on genotype data imputed from SNPs present on the

intersection of these platforms (n=730,803 SNPs). Before imputation, data were cleaned using the quality control (QC) pipeline developed by the eMERGE Genomics Working Group(30). Genomic imputation was then performed for samples using IMPUTE2(31) and data from the 1000 genomes project(32). Imputed SNPs were retained if their call probability was $\geq 90\%$. A total of 7,604,479 SNPs with mean call rate 99.0% were included in the imputed SNP dataset.

Samples were classified as being of European or African descent using ancestry informative markers (AIMs) input into STRUCTURE using Hapmap reference populations(33). Principal components (PC) analysis was also performed on AIMs using the EIGNESTRAT method implemented in PLINK(34, 35). To minimize confounding by population stratification, GWAS analyses were adjusted for PCs 1 and 2. Sensitivity analysis was performed in subjects with greater than 90% STRUCTURE-defined European ancestry. Classical HLA alleles were imputing from SNP data on genome-wide platforms using SNP2HLA with the Type 1 Diabetes Genetics Consortium (T1DGC) reference panel and a marker window size of 1000(36). Additional information regarding imputation, quality control, functional SNP analysis, and analysis of race/ethnicity is available in the Supplemental Materials.

Replication of GWAS SNPs—Genotyping and quality control procedures in the replication cohort used a polymerase chain reaction-high resolution melting procedure and are described in detail in the Supplemental Materials. MAFs and the distribution of genotype groups were compared in HIT patients, Ab^{pos} and Ab^{neg} patients.

SNP Association with PF4/heparin Antibody Development—The SHIP samples were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0. Hybridization of genomic DNA was done in accordance with the manufacturer's standard recommendations. Genotypes were determined using the Birdseed2 clustering algorithm. For QC purposes, several control samples were added. On the chip level, only subjects with a genotyping rate on QC probe sets (QC call rate) of at least 86% were included. Finally, all arrays had a sample call rate $> 92\%$. The overall genotyping efficiency of the GWA was 98.6 %. Imputation of genotypes in SHIP was performed as described in the discovery cohort with IMPUTE-2 using the 1000Genomes v3 reference panel.

Statistical Analysis

BioVU (Discovery Population)—Differences in characteristics between cases and controls were determined using chi square and t-tests. In GWAS analysis, single-locus tests of association were performed using logistic regression in an additive, dominant, and recessive genetic model adjusted for age, gender, type of heparin anticoagulant (UFH versus LMWH), and PCs 1 and 2 for all SNPs in a total of 951 individuals regardless of race/ethnic group. Significance was considered at a genome-wide-corrected significance level of 5.0×10^{-8} . All statistical analyses were performed using PLINK(35) and SAS 9.3 (SAS, Cary, NC). GWAS SNPs were carried forward for replication based on strength of association, direct genotyping versus imputation, and putative functional consequences.

For the candidate gene analysis, the same statistical models were employed on a restricted set of genotyped SNPs in candidate genes. Candidate genes, including HLA class I and class II genes, were selected based on previous literature and HIT pathophysiology and are listed in Table 1. SNPs were identified from non-imputed data in candidate genes plus or minus 10 kilobases. SNPs in the region between *HLA-A* and *HLA-DPB1* were not included unless located within 10 kilobases of a specified candidate gene. A total of 1,412 SNPs with minor allele frequencies (MAFs) greater than 0.01 were tested and we implemented a Bonferroni cutoff of $\alpha=3.54 \times 10^{-5}$ (0.05/1,412). In an additional exploratory analysis, imputed classical HLA alleles were tested for association with HIT in an unadjusted additive model.

Replication of GWAS SNPs—To test the association of genotyped SNPs and HIT, multivariate logistic regression was used. Age and gender-adjusted odds ratios (OR) and 95% confidence intervals (CI) were generated with HIT defined as the outcome compared to both groups without HIT (Ab^{pos} and Ab^{neg} patients) and compared to Ab^{neg} patients in a recessive model. Results were considered significant with a two-sided $\alpha=0.05$.

SNP Association with PF4/heparin Antibody Development—In the SHIP cohort, we assessed the association between GWAS-associated SNPs and PF4/heparin IgG titer levels (enzyme immunoassay OD levels) as well as formation of positive PF4/heparin antibody tests. Two IgG measurements were available for each patient and mean PF4/heparin IgG ODs were analyzed as continuous variables after square root transformation. Association of SNPs with PF4/heparin IgG levels were determined using linear regression adjusted for age and gender with $\alpha=0.05$. Effect sizes of SNPs are reported using coefficients (β) and standard errors (SEs). To evaluate the association between SNPs (exposure) and anti-PF4/heparin Ab status (dependent variable), conditional logistic regression models with fixed effects were used, adjusting for age and gender. ORs and CIs for positive anti-PF4/heparin status was tested for association with SNP genotype in a recessive model and determined by exact methods with median-unbiased estimates. Statistical analyses were performed in R.

Results

Genome-Wide Association Study in the EMR Discovery Population

A total of 73 HIT cases were identified from BioVU and 67 of these cases were successfully genotyped after QC filters. The mean 4Ts score for HIT cases was 5.2 (standard deviation [SD] 0.8). (Table 2) The clinical characteristics of patients with HIT diagnosis are presented in Supplemental Table S2. We identified 884 controls from BioVU that were matched to HIT cases by exposure to UFH or LMWH and were successfully genotyped. The number of consecutive days of exposure to UFH or LMWH was significantly increased in HIT cases versus controls (9.7 [SD 6.0] versus 3.4 [SD 4.0], $p<0.01$). No other significant differences were observed in baseline characteristics between cases and controls.

Twenty two SNPs on chromosome 14 were significantly associated with HIT at a genome-wide significance level (Figure 1). A summary for the most strongly associated SNPs from the HIT GWAS in the recessive model are presented in Table 3. Nineteen of the twenty two significantly associated SNPs, including the three most strongly associated SNPs, were

imputed SNPs. The three significantly associated, genotyped SNPs were within a 90.7 kb genomic region that included the T Cell-Death Associated Gene 8 (*TDAG8*, also known as G protein coupled receptor [*GPR65*]), the galactosylceramidase gene (*GALC*), and a long inter-genic non-protein coding RNA *LOC283587* (Figure 2).

The most strongly associated SNP in the recessive model was the imputed rs10782473 (OR 18.52 [CI 6.33–54.23] $p=3.18 \times 10^{-9}$) located in *LOC283587*, 60 kb downstream of *TDAG8*. The most strongly associated genotyped SNP was rs1887289 (OR 16.83 [CI 5.90–47.97], $p=1.01 \times 10^{-8}$) located in *LOC283587* 68 kb downstream of *TDAG8*. (Supplementary Figure S2) The SNP rs1887289 was associated with decreased *TDAG8*(*GPR65*) transcription using cis-eQTL studies of healthy individuals according to the Genevar database ($\rho=-0.21$, $p=0.03$) (Supplementary Figure S3). The three significantly associated genotyped SNPs were in high LD ($r^2=0.98$) with each other and moderate LD with the nearby missense SNP rs3742704 ($r^2=0.55$) in Caucasian individuals. (Supplementary Figure S4) The missense *TDAG8* SNP rs3742704 was in moderate LD ($r^2=0.55$) with rs1887289 and nominally associated with HIT (OR 33.48 [CI 8.01–140.00], $p=2.04 \times 10^{-7}$). Rs3742704 codes an isoleucine to leucine change at amino acid 231 in *TDAG8*, but was predicted to be tolerated with a SIFT score of 0.23. No *in silico* functional consequences were observed for any other SNPs significantly associated with HIT. No significantly associated SNPs were observed at genome-wide significance levels in either additive or dominant models.

Similar results were obtained when GWAS analysis was restricted to Caucasian individuals, with twenty five SNPs significantly associated with HIT at genome-wide significance. (Supplementary Figure S5) Rs1887289 was out of HWE in cases but well within HWE in controls ($p=4.39 \times 10^{-5}$ and 0.41 respectively) in Caucasian individuals. Similar HWE statistics were observed for the missense SNP rs3742704 (HWE $p=0.002$ in cases and HWE $p=0.25$ in controls).

Candidate Gene Study in EMR Discovery Population

A summary of the most strongly associated SNPs from the candidate gene study, which were observed in a dominant model, is presented in Table S3 of the Supplemental Materials. When analysis was restricted to those SNPs within candidate genes, fourteen SNPs on chromosome 6 were associated with HIT in a dominant model at $p < 3.54 \times 10^{-5}$. Significantly associated SNPs were located in an 11.4 kb region surrounding the major histocompatibility complex, class II, DR alpha gene (*HLA-DRA*). The strongest association for the candidate gene study was for rs4348358 (MAF=0.38, OR 0.25 [CI 0.15–0.44], $p=2.06 \times 10^{-6}$), located upstream of *HLA-DRA*. (Figure 3) All SNPs within the region displayed a reduction in risk for HIT in carriers of the minor allele. The imputed HLA allele most strongly associated with HIT was *HLA-C*12:03* (OR 3.70 [CI 1.85–7.38], $p=0.0002$), which occurred in 6% of the study population. The most strongly association class II HLA allele was *HLA-DRB1*1301* (OR 2.23 [1.23–4.05], $p=0.008$), which occurred in 7% of the study population.

Association of GWAS SNPs in Independent Population

The two *TDAG8* SNPs rs1887289 and rs3742704 were tested for association with clinical HIT in an independent population, comprised of patients with clinical HIT and those without

clinical HIT who were Ab^{pos} or Ab^{neg}. Rs1887289 (MAF 0.10) was not associated with clinical HIT when compared to all other patients in a recessive model (OR 5.71 [CI 0.47–69.22], $p=0.17$). (Table 4) The missense rs3742704 SNP (MAF 0.09) was not associated with clinical HIT when compared to all other patients in a recessive model (OR 1.90 [CI 0.16–22.76], $p=0.61$). (Table 4) Similar results were obtained when patients with clinical HIT were compared to Ab^{neg} patients in a recessive model (OR 1.031 [CI 0.09–12.46], $p=0.98$). Both SNPs were within HWE (rs3742704 HWE $p=0.82$ and rs1887289 HWE $p=0.53$) and LD between the two SNPs was slightly lower ($r^2=0.43$ and $D'=0.70$) than in the GWAS discovery cohort and the SHIP study.

Association of GWAS SNPs and PF4/heparin Antibody formation

The two *TDAG8* SNPs rs1887289 and rs3742704 were tested for association with PF4/heparin IgG titer levels as well as formation of positive PF4/heparin antibody tests. In a total of 3,781 SHIP participants, rs3742704 was significantly associated with PF4/heparin IgG titer levels in a recessive model (MAF=0.09, $\beta = 0.07$ [SE 0.03], $p=0.03$). (Figure S6) The rs3742704 was also significantly associated with the formation of positive PF4/heparin IgG antibodies in a recessive model (OR 3.09 [CI 1.14–8.13], $p=0.02$). The rs1887289 SNP was not significantly associated with PF4/heparin IgG titer levels or positive PF4/heparin antibody tests. Both SNPs were within HWE (rs3742704 HWE $p=0.77$ and rs1887289 HWE $p=0.41$) and LD between the two SNPs was similar in the SHIP cohort to the GWAS discovery cohort ($r^2=0.61$ and $D' = 0.86$). SNPs did not show any significant associations in additive or dominant models, but only in a recessive model, which was consistent with the GWAS in the discovery population.

Discussion

We performed a GWAS for HIT using an EMR-coupled DNA biobank and attempted to replicate our GWAS associations in an independent population of HIT cases and heparin-treated controls. We observed significant associations with HIT from SNPs near *TDAG8* which were in moderate LD with a missense SNP in *TDAG8*. In a separate population, this missense SNP was also associated with PF4/heparin antibody titer levels and development of positive PF4/heparin antibody tests in non-heparin treated patients. In a replication cohort with clinical HIT, *TDAG8* SNPs only trended towards association with HIT. We also investigated the influence of SNPs from candidate genes on HIT in the discovery cohort. In the candidate gene study, multiple SNPs from the *HLA-DRA* gene were associated with HIT, although these signals were independent of classical HLA alleles. Further study of these SNPs is warranted to assess the potential of *TDAG8* and *HLA-DRA* SNPs in predicting HIT and provide mechanistic insights into HIT pathogenesis.

TDAG8 has a biologically plausible role in HIT pathophysiology, suggesting that variants in this gene may play role in the risk of HIT development. The G-protein coupled receptor *TDAG8* has a role in activation-induced cell death of T cells(37) and attenuates immune-mediated inflammation by regulating the cytokine production of T cells(38, 39). HIT is thought to be a T cell-dependent process, as PF4/heparin complexes have been shown to stimulate T cells in HIT patients(40). Other studies in mice also suggested that PF4/heparin

antibody formation is dependent on T cell involvement, and B cell Ig switching requires helper T cells(41, 42). The SNP rs1887289, which was associated with decreased *TDAG8* expression, or the missense SNP rs3742704, could potentially lead to a decrease in abundance and/or function of the TDAG8 protein. This inhibition of TDAG8 could result in T cell stimulation, which might lead to an increase risk of HIT. However, since the observed number of individuals which are homozygous for these mutations is small, genetic risk factors in *TDAG8* are likely to be helpful predictors in only a minority of patients. Thus the genetic contribution to HIT development may be modest and clinical risk factors may prove to be more important determinants of HIT.

The association of *TDAG8* SNPs with HIT in the discovery cohort as well as PF4/heparin antibodies in SHIP is consistent with studies suggesting a role of T cells in PF4/heparin antibody formation. A recessive genetic model produced significant associations in both cohorts, lending credence to the validity of our results and indicating that two copies of the variant allele are necessary to sufficiently impact gene function and affect HIT risk. However, these results must be interpreted with caution since the association of *TDAG8* SNPs and PF4/heparin antibodies was observed in non-heparin treated patients. The possibility that *TDAG8* SNPs influence HIT development is supported by the biologic plausibility of the role of *TDAG8* and *TDAG8* SNPs in HIT, the trend towards association in the replication cohort, the putative functional effects of the associated *TDAG8* SNPs, and the association of the missense SNP with PF4/heparin antibodies.

In the replication cohort, a lack of statistically significant replication may have been a result of differences in case definition between populations or relatively small sample sizes. HIT cases from the replication cohort were defined with greater rigor and the weaker association in the replication cohort suggests that true effect sizes of variants are smaller than observed in the discovery cohort. The discrepancy in results for the discovery and replication cohorts may also have been due to differences in MAF in HIT cases, which was higher in the discovery cohort than in the replication cohort. In the replication cohort, the MAF for both SNPs was higher in HIT cases than in controls, although the absolute number of homozygous variants was very small, which limited our ability to detect a positive association in the replication cohort. Therefore, the lack of statistically significant association in the replication cohort may have been due to small sample size. In addition, differences in LD and MAF underscore the possibility of underlying genetic differences between populations that may have contributed to differences in SNP associations.

The candidate gene study in the discovery cohort also identified *HLA-DRA* alleles as potential predictors of HIT. The *HLA-DR* region provides a particularly attractive candidate gene for HIT, since SNPs in this region have been previously associated with multiple immune-mediated adverse drug reactions(14–16) and class II HLA molecules are directly involved in adaptive immune response resulting in antibody formation. PF4 has been shown to promote monocyte survival and induce monocyte differentiation into macrophages that lack the HLA-DR antigen(43). Expression of the HLA-DR3 antigen has been previously associated with PF4/heparin antibody development after cardiopulmonary bypass(44), although this does not directly support our observed *HLA-DRA* SNP association since the SNP cannot be directly extrapolated to the HLA DR3 serotype and the direction of effect of

the SNP is different from that of the DR3 serotype. Since *HLA-DRA* is not extensively polymorphic and no imputed *HLA-DRA* alleles were strongly associated with HIT from our study, any effect of the observed *HLA-DRA* SNP association on HIT pathogenesis appears to be independent of classical HLA alleles.

Several previous investigations of genetic risk factors for HIT have focused on identifying polymorphisms in genes coding for Fc receptor subunits, which are expressed on platelets and stimulated by PF4/heparin immune complexes(9–13). The Fc receptor RIIA (*FCGR2A*) –H131R and Fc receptor RIIIA (*FCGR3A*) –F158V polymorphisms have been the most extensively investigated. Our study investigated multiple SNPs within *FCGR2A* in the discovery cohort, including the –H131R polymorphism, and we did not observe any significant associations. However, the *FCGR3A* gene was not well covered in our cohort and the –F158V polymorphism was not included in our study. Our results suggest that the *FCGR2A* –H131R polymorphism is not a genetic risk factor for HIT, but the association of the *FCGR3A* –F158V polymorphism with HIT warrants further investigation.

Several limitations are worthy of mention in this study. Firstly, the number of HIT cases in the discovery cohort was small, which may have limited our power to detect an initial association in both the GWAS and candidate gene studies. However, multiple significant SNP effects were observed for both GWAS and candidate gene studies, indicating that this number of cases was sufficient to identify the observed associations. In this study, we were able to detect only common variants associated with HIT and any rare or uncommon variants with modest effects on HIT pathogenesis would not be identified using our approach. The HIT cases were also retrospectively identified in an EMR, which precluded the acquisition of confirmatory functional assay testing for HIT. Although our HIT cases were independently reviewed by multiple investigators, we cannot exclude the possibility of classification bias in our observational HIT case and control cohort. Although some misclassification is likely among HIT cases, the use of a high OD cutoff (0.7 AU) for PF4/heparin antibody testing and the prevalence in known HIT risk factors among cases, such as female sex, surgery, durations of exposure over 6 days, high 4Ts scores, and increased age suggest accurate classification among our cohort. Finally, the duration of UFH or LMWH treatment in our discovery cohort was higher in cases than in controls and many controls may not have been treated for a sufficient time to develop HIT.

In summary, we describe the first GWAS for HIT using an EMR coupled to a DNA biobank. Several SNPs in *TDAG8*, a biologically plausible gene regulating T cell function, were significantly associated with HIT in the discovery population and a missense SNP was associated with formation of PF4/heparin antibodies in a non-heparin treated population. *TDAG8* SNPs also trended towards association with HIT in an independent replication cohort. In a candidate gene study, several SNPs in *HLA-DRA* were associated with HIT, although these signals were independent of classical HLA alleles.. Further study of *TDAG8* and *HLA-DRA* SNPs is warranted to assess their potential in predicting HIT and provide mechanistic insights into HIT pathogenesis. These studies could include additional genetic association and functional studies, as well as high resolution HLA typing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

JHK is supported by the VUMC Clinical Pharmacology Training grant (T32 GM07569). BioVU is supported by institutional funding and by the Vanderbilt CTSA grant UL1 TR000445 from NCATS/NIH. A portion of the genome-wide genotyping was funded by NIH grants RC2GM092618 from NIGMS/OD, U01HG006378 from NHGRI/NIGMS, and U19HL065962 from NHGRI/NIGMS. SLV is supported by an early career award from the PhRMA foundation.

SHIP is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, 03ZIK012, and 03IS2061A), the Ministry of Cultural Affairs, the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED). Genome-wide data were also supported by a joint grant from Siemens Healthcare, Erlangen, Germany. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG and the Caché Campus program of the InterSystems GmbH.

The part of the study performed in Tours (France) with the independent replication population was supported by the IRTH (Institut pour la Recherche sur la Thrombose et l'Hémostase).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Warkentin TE, Levine MN, Hirsh J, et al. Heparin-induced thrombocytopenia in patients treated with low-molecular-weight heparin or unfractionated heparin. *N Engl J Med*. 1995; 332(20):1330–1335. [PubMed: 7715641]
2. Girolami B, Prandoni P, Stefani PM, et al. The incidence of heparin-induced thrombocytopenia in hospitalized medical patients treated with subcutaneous unfractionated heparin: a prospective cohort study. *Blood*. 2003; 101(8):2955–2959. [PubMed: 12480713]
3. Greinacher A, Farner B, Kroll H, et al. Clinical features of heparin-induced thrombocytopenia including risk factors for thrombosis. A retrospective analysis of 408 patients. *Thromb Haemost*. 2005; 94(1):132–135. [PubMed: 16113796]
4. Franchini M. Heparin-induced thrombocytopenia: an update. *Thromb J*. 2005; 3:14. [PubMed: 16202170]
5. Martel N, Lee J, Wells PS. Risk for heparin-induced thrombocytopenia with unfractionated and low-molecular-weight heparin thromboprophylaxis: a meta-analysis. *Blood*. 2005; 106(8):2710–2715. [PubMed: 15985543]
6. Linkins LA, Dans AL, Moores LK, et al. Treatment and prevention of heparin-induced thrombocytopenia: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012; 141(2 Suppl):e495S–e530S. [PubMed: 22315270]
7. Watson H, Davidson S, Keeling D. Guidelines on the diagnosis and management of heparin-induced thrombocytopenia: second edition. *Br J Haematol*. 2012; 159(5):528–540. [PubMed: 23043677]
8. Cuker A. Heparin-induced thrombocytopenia: present and future. *J Thromb Thrombolysis*. 2011; 31(3):353–366. [PubMed: 21327506]
9. Pamela S, Anna Maria L, Elena D, et al. Heparin-induced thrombocytopenia: the role of platelets genetic polymorphisms. *Platelets*. 2013; 24(5):362–368. [PubMed: 22793995]
10. Rollin J, Pouplard C, Leroux D, et al. Impact of polymorphisms affecting the ACP1 gene on levels of antibodies against platelet factor 4-heparin complexes. *J Thromb Haemost*. 2013; 11(8):1609–1611. [PubMed: 23621699]
11. Rollin J, Pouplard C, Gratacap MP, et al. Polymorphisms of protein tyrosine phosphatase CD148 influence FcγRIIA-dependent platelet activation and the risk of heparin-induced thrombocytopenia. *Blood*. 2012; 120(6):1309–1316. [PubMed: 22677127]

12. Poupard C, Cornillet-Lefebvre P, Attaoua R, et al. Interleukin-10 promoter microsatellite polymorphisms influence the immune response to heparin and the risk of heparin-induced thrombocytopenia. *Thromb Res.* 2012; 129(4):465–469. [PubMed: 22239992]
13. Gruel Y, Poupard C, Lasne D, et al. The homozygous FcγRIIIa-158V genotype is a risk factor for heparin-induced thrombocytopenia in patients with antibodies to heparin-platelet factor 4 complexes. *Blood.* 2004; 104(9):2791–2793. [PubMed: 15191947]
14. Daly AK, Donaldson PT, Bhatnagar P, et al. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. *Nat Genet.* 2009; 41(7):816–819. [PubMed: 19483685]
15. Mallal S, Phillips E, Carosi G, et al. HLA-B*5701 screening for hypersensitivity to abacavir. *N Engl J Med.* 2008; 358(6):568–579. [PubMed: 18256392]
16. Phillips EJ, Mallal SA. HLA-B*1502 screening and toxic effects of carbamazepine. *N Engl J Med.* 2011; 365(7):672. author reply 3. [PubMed: 21848472]
17. Selleng S, Greinacher A, Ittermann T, et al. Heparin/PF4 antibodies formation after heparin treatment: temporal aspects and long-term follow-up. *Am Heart J.* 2009; 158(1):e15. [PubMed: 19540379]
18. Greinacher A, Kohlmann T, Strobel U, et al. The temporal profile of the anti-PF4/heparin immune response. *Blood.* 2009; 113(20):4970–4976. [PubMed: 19109231]
19. Warkentin TE, Sheppard JA, Moore JC, et al. Studies of the immune response in heparin-induced thrombocytopenia. *Blood.* 2009; 113(20):4963–4969. [PubMed: 19144981]
20. Warkentin TE, Kelton JG. Temporal aspects of heparin-induced thrombocytopenia. *N Engl J Med.* 2001; 344(17):1286–1292. [PubMed: 11320387]
21. Roden DM, Pulley JM, Basford MA, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther.* 2008; 84(3):362–369. [PubMed: 18500243]
22. Denny JC, Spickard A 3rd, Johnson KB, et al. Evaluation of a method to identify and categorize section headers in clinical documents. *J Am Med Inform Assoc.* 2009; 16(6):806–815. [PubMed: 19717800]
23. Denny JC, Miller RA, Waitman LR, et al. Identifying QT prolongation from ECG impressions using a general-purpose Natural Language Processor. *Int J Med Inform.* 2009; 78(Suppl 1):S34–S42. [PubMed: 18938105]
24. Cuker A, Gimotty PA, Crowther MA, et al. Predictive value of the 4Ts scoring system for heparin-induced thrombocytopenia: a systematic review and meta-analysis. *Blood.* 2012; 120(20):4160–4167. [PubMed: 22990018]
25. Volzke H, Alte D, Schmidt CO, et al. Cohort profile: the study of health in Pomerania. *Int J Epidemiol.* 2011; 40(2):294–307. [PubMed: 20167617]
26. Greinacher A, Holtfreter B, Krauel K, et al. Association of natural anti-platelet factor 4/heparin antibodies with periodontal disease. *Blood.* 2011; 118(5):1395–1401. [PubMed: 21659541]
27. Ludemann J, Piek M, Wood WG, et al. [Methods for quality assurance of medical examination in epidemiological field studies: the "Study of Health in Pomerania" (SHIP)]. *Gesundheitswesen.* 2000; 62(4):234–243. [PubMed: 10844821]
28. Juhl D, Eichler P, Lubenow N, et al. Incidence and clinical significance of anti-PF4/heparin antibodies of the IgG, IgM, and IgA class in 755 consecutive patient samples referred for diagnostic testing for heparin-induced thrombocytopenia. *Eur J Haematol.* 2006; 76(5):420–426. [PubMed: 16466367]
29. Althaus K, Strobel U, Warkentin TE, et al. Combined use of the high heparin step and optical density to optimize diagnostic sensitivity and specificity of an anti-PF4/heparin enzyme-immunoassay. *Thromb Res.* 2011; 128(3):256–260. [PubMed: 21620439]
30. Zuvich RL, Armstrong LL, Bielinski SJ, et al. Pitfalls of merging GWAS data: lessons learned in the eMERGE network and quality control procedures to maintain high data quality. *Genet Epidemiol.* 2011; 35(8):887–898. [PubMed: 22125226]
31. Howie B, Fuchsberger C, Stephens M, et al. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet.* 2012; 44(8):955–959. [PubMed: 22820512]

32. Abecasis GR, Altshuler D, Auton A, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010; 467(7319):1061–1073. [PubMed: 20981092]
33. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000; 155(2):945–959. [PubMed: 10835412]
34. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006; 38(8):904–909. [PubMed: 16862161]
35. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3):559–575. [PubMed: 17701901]
36. Jia X, Han B, Onengut-Gumuscu S, et al. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One*. 2013; 8(6):e64683. [PubMed: 23762245]
37. Choi JW, Lee SY, Choi Y. Identification of a putative G protein-coupled receptor induced during activation-induced apoptosis of T cells. *Cell Immunol*. 1996; 168(1):78–84. [PubMed: 8599842]
38. Onozawa Y, Fujita Y, Kuwabara H, et al. Activation of T cell death-associated gene 8 regulates the cytokine production of T cells and macrophages in vitro. *Eur J Pharmacol*. 2012; 683(1–3):325–331. [PubMed: 22445881]
39. Onozawa Y, Komai T, Oda T. Activation of T cell death-associated gene 8 attenuates inflammation by negatively regulating the function of inflammatory cells. *Eur J Pharmacol*. 2011; 654(3):315–319. [PubMed: 21238451]
40. Bacsí S, De Palma R, Visentin GP, et al. Complexes of heparin and platelet factor 4 specifically stimulate T cells from patients with heparin-induced thrombocytopenia/thrombosis. *Blood*. 1999; 94(1):208–215. [PubMed: 10381515]
41. Fleischer J, Grage-Griebenow E, Kasper B, et al. Platelet factor 4 inhibits proliferation and cytokine release of activated human T cells. *J Immunol*. 2002; 169(2):770–777. [PubMed: 12097379]
42. Suvarna S, Rauova L, McCracken EK, et al. PF4/heparin complexes are T cell-dependent antigens. *Blood*. 2005; 106(3):929–931. [PubMed: 15845897]
43. Scheuerer B, Ernst M, Durrbaum-Landmann I, et al. The CXC-chemokine platelet factor 4 promotes monocyte survival and induces monocyte differentiation into macrophages. *Blood*. 2000; 95(4):1158–1166. [PubMed: 10666185]
44. Paparella D, Micelli M, Favoino B, et al. Anti-heparin-platelet factor 4 antibodies after cardiopulmonary bypass: role of HLA expression. *Haematologica*. 2001; 86(3):326–327. [PubMed: 11255285]
45. Harris K, Nguyen P, Van Cott EM. Platelet PIA2 Polymorphism and the risk for thrombosis in heparin-induced thrombocytopenia. *Am J Clin Pathol*. 2008; 129(2):282–286. [PubMed: 18208809]
46. Carlsson LE, Lubenow N, Blumentritt C, et al. Platelet receptor and clotting factor polymorphisms as genetic risk factors for thromboembolic complications in heparin-induced thrombocytopenia. *Pharmacogenetics*. 2003; 13(5):253–258. [PubMed: 12724616]

What is known on this topic	<ul style="list-style-type: none">• Heparin-induced thrombocytopenia (HIT) is an unpredictable, potentially catastrophic adverse effect of heparin treatment• The inability to predict HIT is a critical problem with heparin administration and a priori identification of patients with a high HIT risk would allow for prevention of HIT.• Although few candidate gene studies have identified genetic polymorphisms associated with HIT, none of these studies have identified genetic risk factors that can reliably predict HIT.
What this paper adds	<ul style="list-style-type: none">• We describe the first genome-wide association study and candidate gene study for heparin-induced thrombocytopenia (HIT).• SNPs in the <i>TDAG8</i> and <i>HLA-DRA</i> genes were implicated as possible predictors of HIT.• Further study of <i>TDAG8</i> and <i>HLA-DRA</i> SNPs is warranted to assess their potential in predicting HIT and provide mechanistic insights into HIT pathogenesis.

Figure 1.

Manhattan plot of genotyped SNPs associated with heparin-induced thrombocytopenia in a recessive model, adjusted for age, gender, UFH versus LMWH, and principal components 1 and 2. P values on the $-\log_{10}$ scale are plotted on the vertical axis and the chromosomal position is plotted along the horizontal axis. Dots representing SNPs on different chromosomes have alternating orange and blue color. The red line indicates the genome-wide significance threshold of $\alpha=5 \times 10^{-8}$.

Figure 2.

Locuszoom plot of most strongly associated SNPs in *TDAG8* (*GPR65*) and the surrounding region using imputed SNPs. P values were generated using logistic regression adjusted for age, gender, UFH versus LMWH, and principal components 1 and 2 in a recessive model. P values on the $-\log_{10}$ scale are plotted on the left vertical axis and the chromosomal position is plotted along the horizontal axis along with the gene names and size of flanking region. The most strongly associated SNP rs10782473 is indicated by a purple diamond and pairwise LD (r^2) with this SNP is indicated by dot color as described in the legend in the upper right corner. The right vertical axis indicates the regional recombination rate (cM/Mb) which is overlaid in light blue.

Figure 3.

Locuszoom plot of most strongly associated SNPs from the candidate gene study in *HLA-DRA* and the surrounding region plus or minus 10,000 base pairs. P values were generated using logistic regression adjusted for age, gender, UFH versus LMWH, and principal components 1 and 2 in a recessive model. P values on the $-\log_{10}$ scale are plotted on the left vertical axis and the chromosomal position is plotted along the horizontal axis along with the gene names and size of flanking region. Pairwise LD (r^2) is indicated by dot color as described in the legend in the upper left corner. The right vertical axis indicates the regional recombination rate (cM/Mb) which is overlaid in light blue.

Table 1

List of candidate genes with description and rationale for inclusion in study.

Gene	Number of SNPs within gene ¹	Description	Rationale
<i>FCGR2A, FCGR3A, FCGR3B</i>	12, 1, 1	Fc fragment of IgG, low affinity IIa/IIa/IIIb receptor	Previous association(13); central role in HIT pathophysiology
<i>PF4</i>	20	platelet factor 4	Central role in HIT pathophysiology
<i>ITGA2B</i>	22	integrin, alpha 2b (platelet glycoprotein IIb)	Previous association(45, 46)
<i>PECAM1</i>	35	platelet/endothelial cell adhesion molecule 1	Previous association(9)
<i>IL10</i>	38	interleukin 10	Previous association(12)
<i>PTPRJ</i>	101	protein tyrosine phosphatase, receptor type, J	Previous association(11)
<i>ACPI</i>	23	acid phosphatase 1, soluble	Previous association(10)
<i>HLA-A, HLA-B, HLA-C, HLA-DR, HLA-DP, HLA-DQ</i>	79, 151, 178, 311, 736, 419	class I and II human leukocyte antigen genes	Previous <i>HLA-DR</i> association;(43, 44) immune-mediated adverse drug reaction requires antigen presentation

HIT indicates heparin-induced thrombocytopenia; HLA, human leukocyte antigen; SNP, single nucleotide polymorphism.

¹) Indicates SNPs of any minor allele frequency within 10,000 base pairs of 3' and 5' end of gene.

Table 2

Baseline characteristics in cases versus controls of genotyped samples for GWAS analysis after QC filters.

Characteristic ¹	Cases (n=67)	Controls (n=884)	p value ²
Age (years), mean (SD)	56.0 (16.6)	55.4 (±15.8)	0.76
Gender (female)	34 (51)	403 (46)	0.41
Height (cm), mean (SD)	168 (18)	170 (11)	0.28
Weight (kg), mean (SD)	87.6 (26.2)	87.4 (25.0)	0.94
BMI (kg/m ²), mean (SD)	29.5 (8.9)	30.0 (10.6)	0.74
Heparin agent			0.74
Heparin	59 (88)	734 (83)	
Enoxaparin	9 (13)	150 (17)	
Heparin agent duration (days), mean (SD) ³	9.7 (6.0)	3.4 (4.0)	<0.01
Race/ethnicity ⁴			0.90
White	57 (85)	754 (85)	
Black	8 (12)	108 (12)	
Unknown	1 (1)	6 (1)	
Hispanic	0 (0)	8 (1)	
Asian	0	1 (0.1)	
Other	1 (1%)	6 (1)	
Baseline platelets (×10 ³ /μL), mean (SD)	330 (116)	332 (144)	0.92
4Ts Score, mean (SD) ⁵	5.2 (0.8)	-	-

BMI indicates body mass index; cm, centimeters; kg, kilograms; SD, standard deviation.

¹) Values represent number (percentage) unless otherwise specified²) P values represent results of chi square and t-tests where appropriate³) Longest stretch of consecutive days of exposure to heparin agent⁴) Race/ethnicity values are those recorded in the EMR⁵) 4Ts scored according to American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (Linkins et al. Chest 2012;141:e495S–530S)

Table 3

Top associated genotyped SNPs in a recessive model.

Chr	SNP	Gene ¹	Nearest gene ²	Function ¹	Risk allele	MAF (Cases)	MAF (Controls)	OR (CI) ³	p value ³
Imputed SNPs									
14	rs10782473	<i>LOC283587</i>	<i>TDAG8 (GPR65)</i>	intron	G	0.22	0.11	18.52 (6.33–54.23)	3.18×10 ^{−9}
14	rs7161383	<i>LOC283587</i>	<i>TDAG8 (GPR65)</i>	intron	C	0.22	0.11	18.48 (6.31–54.10)	3.26×10 ^{−9}
14	rs4464016	intergenic	<i>LOC283587</i>	intron	C	0.21	0.11	16.73 (5.87–47.67)	3.33×10 ^{−9}
Genotyped SNPs									
14	rs1887289	<i>LOC283587</i>	<i>TDAG8 (GPR65)</i>	intron	T	0.21	0.11	16.83 (5.90–47.97)	1.01 × 10 ^{−8}
14	rs8014798	<i>GALC</i>	<i>TDAG8 (GPR65)</i>	intron	A	0.20	0.10	20.59 (5.97–70.99)	1.27 × 10 ^{−8}
14	rs11159833	<i>GPR65</i>	<i>GALC</i>	intron	T	0.22	0.11	12.90 (4.23–39.33)	2.30 × 10 ^{−8}
14	rs3742704	<i>GPR65</i>	<i>LOC283587</i>	Missense (Ile-Leu)	G	0.19	0.09	33.48 (8.01–140.00)	2.04 × 10 ^{−7}

Chr indicates chromosome; CI, 95% confidence interval; MAF, minor allele frequency; OR, odds ratio; SNP, single nucleotide polymorphism;

- 1) Gene represents the gene in which a SNP directly resides, if any, and function indicates the putative functional impact of a SNP in this gene, as predicted by SNP Nexus.
- 2) Nearest gene represents nearest gene outside of the gene in which the SNP resides, if any.
- 3) Odds ratios, 95% confidence intervals, and p values were generated in a recessive model adjusted for age, gender, heparin type (UFH versus LMWH), and principal components 1 and 2.

Table 4
Association and genotype distribution of *TDAG8* SNPs in the replication population

SNP	Genotype distribution ¹	MAF	HWE p value ²	OR (95%CI) ³	p value ³
rs1887289					
Overall	345 / 78 / 3	0.10	0.53	5.71 (0.47–69.22)	0.17
Ab ^{neg}	144 / 34 / 0	0.10	0.16		
Ab ^{pos}	127 / 26 / 1	0.09	0.79		
HIT	74 / 18 / 2	0.12	0.48		
rs3742704					
Overall	353 / 70 / 3	0.09	0.82	1.90 (0.16–22.76)	0.61
Ab ^{neg}	148 / 28 / 2	0.09	0.61		
Ab ^{pos}	129 / 25 / 0	0.08	0.27		
HIT	76 / 17 / 1	0.10	0.96		

Ab indicates antibody; HIT, heparin-induced thrombocytopenia; HWE, Hardy Weinberg Equilibrium; MAF, minor allele frequency; OR, odds ratio; SNP, single nucleotide polymorphism; 95% CI, 95% confidence interval.

¹) Genotype distributions represent number of individuals with homozygous wild type, heterozygous, and homozygous variant genotypes, in that order.

²) Hardy Weinberg Equilibrium p values are generated using chi square tests.

³) Odds ratios, confidence intervals and p values are adjusted for age and gender in a recessive model when comparing patients with clinical HIT to all other patients (AB^{pos} and AB^{neg}).