ORIGINAL ARTICLE

Genome-wide studies of von Willebrand factor propeptide identify loci contributing to variation in propeptide levels and von Willebrand factor clearance

A. B. OZEL,* B. MCGEE,† D. SIEMIENIAK,† P. M. JACOBI,‡ S. L. HABERICHTER,‡ L. C. BRODY,§
J. L. MILLS,¶ A. M. MOLLOY,** D. GINSBURG,*††‡‡ J. Z. LI* and K. C. DESCH‡‡
*Human Genetics, University of Michigan, ; †Howard Hughes Medical Institute, University of Michigan, Ann Arbor, MI; ‡The Blood Center of Wisconsin, Milwaukee, WI; § National Human Genome Research Institute, National Institutes of Health; ¶Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA; **School of Medicine, Trinity College Dublin, Dublin, UK; ††Department of Internal Medicine, University of Michigan; and ‡‡Department of Pediatrics and Communicable Disease, University of Michigan, Ann Arbor, MI, USA

Essentials

- Variants at *ABO*, von Willebrand Factor (*VWF*) and 2q12 contribute to the variation in plasma in VWF.
- We performed a genome-wide association study of plasma VWF propeptide in 3,238 individuals.
- *ABO*, *VWF* and 2q12 loci had weak or no association or linkage with plasma VWFpp levels.
- VWF associated variants at *ABO*, *VWF* and 2q12 loci primarily affect VWF clearance rates.

Summary. *Background:* Previous studies identified common variants at the *ABO* and *VWF* loci and unknown variants in a chromosome 2q12 linkage interval that contributed to the variation in plasma von Willebrand factor (VWF) levels. Whereas the association with *ABO* haplotypes can be explained by differential VWF clearance, little is known about the mechanisms underlying the association with *VWF* single-nucleotide polymorphisms (SNPs) or with variants in the chromosome 2 linkage interval. VWF propeptide (VWFpp) and mature VWF are encoded by the *VWF* gene and secreted at the same rate, but have different plasma half-lives. Therefore, comparison of VWFpp and VWF association signals can be used to assess whether the variants are primarily affecting synthesis/secretion or clearance. *Methods:* We measured plasma VWFpp levels and performed genome-wide linkage and association studies in 3238 young and healthy individuals for whom VWF levels had been analyzed previously. *Results and conclusions:* Common variants in an intergenic region on chromosome 7q11 were associated with VWFpp levels. We found that *ABO* serotype-specific SNPs were associated with VWFpp levels in the same direction as for VWF, but with a much lower effect size. Neither the association at *VWF* nor the linkage on chromosome 2 previously reported for VWF was observed for VWFpp. Taken together, these results suggest that the major genetic factors affecting plasma VWF levels, i.e. variants at *ABO*, *VWF* and a locus on chromosome 2, operate primarily through their effects on VWF clearance.

Keywords: genetic linkage analysis; genome-wide association study; venous thromboembolism; von Willebrand disease; von Willebrand factor.

Correspondence: Karl C. Desch, Department of Pediatrics and Communicable Disease, University of Michigan, Life Sciences Institute, Rm 5028, 210 Washtenaw Ave, Ann Arbor, MI 48109, USA.
Tel: +1 734 647 4784; fax: +1 734 936 2888.
E-mail: kdesch@umich.edu

Introduction

Plasma von Willebrand factor (VWF) levels vary approximately five-fold among healthy individuals, and are highly heritable [1–5]. Individuals with VWF levels at the extremes of the normal distribution are at risk for common disorders of hemostasis: low levels cause the common bleeding disorder type I von Willebrand disease (VWD); and high levels are associated with increased risks for both venous and arterial thrombosis [6,7]. Therefore, identification of the genetic factors affecting VWF levels may lead to improved, personalized care for

patients with VWD, arterial thrombosis, and venous thromboembolic disease [8].

Previous genome-wide association studies (GWASs) demonstrated that common variants at *ABO*, *VWF* and other loci explained ~ 12% of the variation in plasma VWF levels [9]. More recently, we reported genome-wide linkage and association analyses of plasma VWF levels in a cohort of young healthy individuals [3], and confirmed the association between *ABO* A1 and B alleles with elevated VWF levels relative to the O and A2 alleles. At *VWF*, haplotypes containing common variants near the factor VIII-binding (D′ domain) and propeptide-coding (D1 and D2 domains) regions were also significantly associated with VWF levels [3,10]. Using the sibling structure in our cohorts, we performed linkage analysis, and identified an interval on chromosome (Chr) 2q12 explaining ~ 19% of the variation in VWF levels [3]. Although the effect may not be direct [11], the major allelic variants at *ABO* influence VWF levels through altered VWF clearance [12–14], whereas the basic mechanisms underlying the effect of the variants in *VWF* and in Chr 2q12 remain unknown.

A specialized post-translational modification system facilitates the storage and secretion of VWF multimers [15]. The VWF propeptide (VWFpp) is cleaved from pro-VWF dimers, but remains non-covalently bound to the developing VWF multimers until both are secreted into flowing blood. VWFpp is rapidly cleared from the circulation, with a half-life of 2–3 h, whereas multimeric VWF has a longer half-life of 8–12 h [16]. As alterations in VWF synthesis or secretion would be expected to affect VWF and VWFpp levels similarly, relative differences in the steady-state levels of these two proteins in plasma should reflect differences in their clearance rates. For this reason, elevation in the VWFpp/VWF ratio has been used to differentiate the subset of individuals with VWD resulting from rapid clearance of VWF from those with reduced synthesis or secretion [17,18].

To identify genetic variants affecting plasma VWFpp levels and to determine whether single-nucleotide polymorphisms (SNPs) previously associated with VWF variation operate primarily through altered synthesis/secretion or clearance mechanisms, we measured VWFpp levels in plasma samples from the Genes and Blood Clotting Study (GABC) and the Trinity Student Study (TSS), in which VWF levels had already been measured and their association results reported [3]. Using an initial measure of VWFpp in the GABC cohort, we identified a strong association with a non-synonymous SNP in the VWFpp-coding portion of *VWF*. However, this association was not present when a different set of mAbs were used to determine VWFpp levels, suggesting that an altered epitope resulting from this SNP had generated the apparent association. We therefore used only VWFpp assay results without interference from this altered epitope in association and linkage studies, and identified significant associations between VWFpp and variants at *ABO* and a new

locus on Chr 7. Comparison of these results with those of our previously reported VWF studies [3] suggest that variants at *VWF* and the Chr 2 linkage interval modify VWF levels mainly through clearance mechanisms.

Materials and methods

GABC

One thousand one hundred and eighty-nine healthy siblings between the ages of 14 years and 35 years were recruited from the students and staff at the University of Michigan, Ann Arbor between June 2006 and January 2009 [3,19]. Subjects who reported that they were pregnant or had a known bleeding disorder or chronic illness requiring regular medical care were excluded. All subjects gave informed consent prior to their participation [20].

TSS

Healthy Irish individuals aged 18–28 years, attending Trinity College of the University of Dublin, were recruited between 2003 and 2004 for genetic analysis of nutrition-related and diet-related traits. A total of 2490 participants completed questionnaires and donated blood samples. Ethical approval from the Dublin Federated Hospitals Research Ethics Committee was obtained and reviewed by the Office of Human Subjects Research at the United States National Institute of Health. Participants provided written informed consent prior to recruitment.

Plasma VWFpp levels

VWFpp levels were measured in the GABC and TSS cohorts with AlphaLISA (Perkin-Elmer, Waltham, MA, USA) [3]. To create VWFpp-specific AlphaLISA assays, mAbs were produced in the Blood Center of Wisconsin Hybridoma Core Laboratory (Milwaukee, WI, USA) by immunizing mice with purified recombinant human VWFpp. mAbs were purified from ascites, and epitope mapping was performed by the use of standard ELISAs, with truncated fragments of recombinant VWFpp as targets (Fig. S1). For TSS plasma samples, which were collected with EDTA as an anticoagulant, 0.8 mM $CaCl_2$ was added to the assay buffer to correct for signal loss associated with the presence of EDTA (Fig. S2).

VWFpp levels were calculated on the assumption of a 1 : 1 ratio of VWF to VWFpp in laboratory control plasma from pooled donors (FACT; George King Biomedical, Overland Park, KS, USA). Each sample was independently assayed at least four times. The mean sample coefficients of variation were 1.0% (GABC, anti-D2), 1.1% (GABC, anti-D1), and 2.8% (TSS, anti-D1). After quality control, 3381 subjects (1110 GABC; 2271 TSS) had VWFpp levels available for heritability analysis, and 3238 had them available for association and linkage testing.

Phenotype data processing

The raw VWFpp distribution was normalized by log transformation, and adjusted for the effects of age, gender, and population genetic structure. We evaluated the correlation between each of the first 10 principal component (PC) scores and the age-corrected and gender-corrected VWFpp levels for GABC and TSS separately, and used Pearson's correlation coefficients and *P*-values to determine which PCs have a significant impact ($P < 0.05$) on the phenotype. After analyzing Pearson's correlation coefficients, *P*-values and genomic control (GC) factors, we selected the PC(s) that were highly correlated with the two traits and resulted in the lowest GC factors. For use in GWAS and linkage studies, log-transformed VWFpp levels were adjusted for age, gender, and the selected PC (s): PC2 and PC7 for GABC, and PC4 for TSS.

Genotyping, imputation, and quality control

GABC: Details of the genotyping and data-cleaning process have been previously published [3,20,21]. The final cleaned dataset contained 763 195 SNPs and 1152 subjects representing 489 sibships. Imputation was carried out by the GENEVA Consortium Data Coordinating Center, using BEAGLE v3.3.1 [22] on a set of 767 243 genotyped SNPs. The final dataset included ~ 7.50 million SNPs. We then removed SNPs with low imputation quality ($R^2 < 0.3$) and low allele frequency (minor allele frequency [MAF] of < 2%), resulting in ~ 5.95 million SNPs.

TSS: After extensive cleaning, the final dataset contained 757 577 SNPs.

For imputation, we first pre-phased the cleaned dataset by using SHAPEIT2 v2.r778 [23], and then used IMPUTE2 v2.3.0 [24] and the 1000 Genomes Phase I release v3 integrated haplotypes (produced by the use of SHAPEIT2 in December 2013) for imputation. This generated 10 520 121 imputed SNPs. After removal of the SNPs with low imputation quality ($R^2 < 0.3$), low allele frequency (MAF of < 2%), and low call rate (< 95%), and that failed the test of Hardy–Weinberg equilibrium ($P < 1.0 \times 10^{-6}$), the final dataset contained ~ 7.37 million SNPs in 2304 individuals.

The TSS and GABC imputed datasets had ~ 4.51 million SNPs in common; this set was used for the meta-analyses. The genome-wide significance level was set at $P = 5.0 \times 10^{-8}$, based on a conservative Bonferroni correction for ~ 1 million independent tests.

Genetic analyses

Population substructure Owing to the presence of sibships of different sizes, GABC samples were analyzed with a two-step approach [25], as detailed previously [3,19].

Association analyses For the GABC cohort, we performed single-SNP association analysis of the transformed and adjusted VWFpp antigen levels by using the imputed set of ~ 5.95 million SNPs. To account for the inferred relatedness and subtle population stratification, we applied a mixed linear model implemented in EMMAX [26]. For the TSS cohort, we performed single-SNP association tests on the set of ~ 7.37 million imputed SNPs by using PLINK (v1.07) [27], assuming an additive mode of allelic effect. For both cohorts, we calculated the GC factor [28] to assess the degree of residual stratification.

Meta-analysis We performed a meta-analysis of GABC and TSS association results with a sample size-weighted approach on the common set of ~ 4.51 million imputed SNPs by using METAL [29]. The GC factor was near 1 for both analyses, i.e. 0.972 and 1.013 for GABC (EMMAX) and TSS (PLINK), respectively, and the association statistics were corrected to reach a GC factor of 1.000 by the use of METAL before being used in the meta-analysis. Regional plots of the significant findings were produced with LOCUSZOOM [30].

Detection of an antibody-specific SNP association with VWFpp VWFpp levels were initially determined by using a pair of monoclonal anti-VWFpp D2 domain antibodies (Fig. S1). In the European subset ($n = 934$) of GABC, a single SNP, rs1800378(T), with a MAF of 33% in our cohort, was associated with VWFpp ($P = 1.15 \times 10^{-12}$, $\beta = 0.051$ IU dL^{-1} per allele in an additive model) (Fig. S3A–C). This SNP encodes a histidine to arginine substitution at position 484 in the D2 domain of VWFpp, and is predicted by POLYPHEN-2 [31] to be a benign variant. It was not significantly associated with plasma VWF levels ($P = 0.52$). We repeated the VWFpp measurement with a second pair of anti-VWFpp D1 domain antibodies (Fig. S1). Although the log-transformed VWFpp levels measured with the anti-D2 domain antibodies were highly correlated with those measured with the anti-D1 domain antibodies (Spearman's $\rho = 0.97$, $P < 0.0001$), VWFpp levels measured with the anti-D1 domain antibodies did not produce a significant signal at rs1800378 ($P = 0.061$; Fig. S3D,E), suggesting that VWFpp bearing this coding change has altered affinity for the anti-D2 domain antibody pair. We used VWFpp levels measured with the anti-D1 domain antibodies in subsequent association and linkage analyses.

Linkage analysis Linkage analysis was carried out with MERLIN-REGRESS [32] on the European sibling subset of GABC and the sibling subset of TSS ($n = 138$). We employed a clustering algorithm and a permutation-based locus-counting approach to calculate empirical *P*-values for the top linkage signals, as previously described [3].

Starting with the MERLIN-REGRESS output, the subsequent analysis was carried out with custom scripts in R [33].

Heritability analysis   For the combined sibling subset of TSS and 1139 GABC individuals (557 sibships), we used two pedigree-based methods for estimating heritability, i.e. intraclass correlation and MERLIN-REGRESS (v1.1.2), as previously described [3,19,32].

Variance explained by association and linkage regions   The Genome-wide Complex Trait Analysis (GCTA) package [34] was used to estimate the proportion of variation in VWFpp levels explained by the entire genome, the top associated SNPs, or intervals representing individual genes or loci, as previously described [3,19].

Haplotype-based association analysis   PLINK [27] was used to carry out haplotype association with the one-degree-of-freedom haplotype-specific test. ABO serotypes (A1, A2, O, and B) were tagged by the three SNPs rs687289, rs8176704, and rs8176749, as reported by Barbalic *et al*. [35].

Results

VWFpp levels are highly heritable

Details of the demographic characteristics of the TSS and GABC cohorts are shown in Table 1. There were differences in the distribution of unadjusted VWFpp levels (Fig. S4A) between the GABC and TSS cohorts (Kolmogorov–Smirnov test, $P = 0.0001$), with median VWFpp levels of 100.0 IU dL^{-1} in GABC and 86.1 IU dL^{-1} in TSS (Table 1). These differences may have resulted from sample collection in acid citrate dextrose for GABC versus in EDTA for TSS that was not adequately corrected for by additions of CaCl$_2$ to the assay buffer (see Materials and methods). However, when the VWFpp values were log-transformed, mean-centered within each cohort, and adjusted for age, sex, and the population structure, the distribution differences between GABC and TSS disappeared (Fig. S4B). On the basis of on interclass correlation among the siblings in the GABC and TSS cohorts, the narrow-sense heritability of VWFpp was 77.6%, consistent with the value of 80.4% obtained with MERLIN-REGRESS. These heritability estimates are similar to those we reported for VWF, i.e. 64.5% and 66.3%, using interclass correlation and MERLIN-REGRESS, respectively [3]. There was a significant positive correlation between the adjusted VWFpp and

Table 1 Characteristics of study participants

Cohort	GABC*	TSS
No. of subjects (*n*)	1152	2304
Age (years) (Q1, Q3)	21 (19, 23)	22 (21, 24)
Female, *n* (%)	721 (63)	1352 (59)
Sibship size (no. of sibships)	1 (13); 2 (366); 3 (94); 4 (22); 5 (5); 6 (2)	2 (66); 3 (2)
Median VWF level (IU dL^{-1}) (Q1, Q3)	100.2 (77.5, 130.7)	99.8 (79.6, 128.1)
Median VWFpp level (IU dL^{-1}) (Q1, Q3)	100.0 (82.2, 122.1)	86.1 (71.8, 104.6)

GABC, Genes and Blood Clotting Study; TSS, Trinity Student Study; VWF, von Willebrand factor; VWFpp, von Willebrand factor propeptide. *Data from previous VWF analysis [3], except for VWFpp level.

Table 2 Top 15 imputed meta-analysis SNPs for VWFpp. Genome-wide significant (*P*-value $<5.0 \times 10^{-8}$) SNPs in GABC (EMMAX) + TSS (Plink) meta-analysis and individual cohorts, sorted by genomic location, with respect to their relationship to the nearest gene

			Meta-analysis†			GABC			TSS		
SNP	Chr	Position*	SNP Allele	*P*-value	Closest gene	Allele freq.	Beta (SE)	*P*-Value	Allele freq.	Beta (SE)	*P*-value
rs10251762	7	68335032	T	2.9×10^{-8}	*AUTS2*	0.078	0.034 (0.011)	2.9×10^{-3}	0.065	0.034 (0.0072)	2.4×10^{-6}
rs12531236	7	68336435	T	4.2×10^{-8}	*AUTS2*	0.081	0.034 (0.010)	1.6×10^{-3}	0.062	0.031 (0.0068)	6.3×10^{-6}
rs10246260	7	68338971	A	2.9×10^{-8}	*AUTS2*	0.078	0.034 (0.011)	2.9×10^{-3}	0.065	0.034 (0.0072)	2.4×10^{-6}
rs55800567	7	68340079	A	2.0×10^{-8}	*AUTS2*	0.079	0.035 (0.011)	2.0×10^{-3}	0.060	0.034 (0.0072)	2.4×10^{-6}
rs10252976	7	68343237	T	3.7×10^{-8}	*AUTS2*	0.92	−0.033 (0.011)	3.9×10^{-3}	0.93	−0.034 (0.0072)	2.4×10^{-6}
rs11977562	7	68349402	A	4.3×10^{-8}	*AUTS2*	0.92	−0.033 (0.011)	3.9×10^{-3}	0.94	−0.034 (0.0072)	2.8×10^{-6}
rs56835261	7	68356258	A	1.6×10^{-8}	*AUTS2*	0.92	−0.035 (0.011)	2.1×10^{-3}	0.93	−0.034 (0.0072)	1.9×10^{-6}
rs8176749	9	136131188	A	1.1×10^{-13}	*ABO*	0.070	0.052 (0.012)	2.0×10^{-5}	0.073	0.040 (0.0066)	8.5×10^{-10}
rs8176746	9	136131322	A	1.0×10^{-13}	*ABO*	0.070	0.052 (0.012)	2.0×10^{-5}	0.073	0.040 (0.0066)	8.3×10^{-10}
rs8176743	9	136131415	A	1.2×10^{-13}	*ABO*	0.070	0.052 (0.012)	2.0×10^{-5}	0.073	0.040 (0.0065)	9.5×10^{-10}
rs8176741	9	136131461	T	1.4×10^{-13}	*ABO*	0.070	0.052 (0.012)	2.0×10^{-5}	0.073	0.040 (0.0066)	1.1×10^{-9}
rs8176725	9	136132617	T	1.9×10^{-9}	*ABO*	0.093	0.042 (0.010)	5.7×10^{-5}	0.099	0.026 (0.0057)	4.5×10^{-6}
rs8176722	9	136132754	T	2.4×10^{-9}	*ABO*	0.093	0.042 (0.010)	7.0×10^{-5}	0.099	0.026 (0.0057)	4.8×10^{-6}
rs687289	9	136137106	T	6.4×10^{-10}	*ABO*	0.36	0.023 (0.0065)	2.7×10^{-4}	0.25	0.020 (0.0039)	4.7×10^{-7}
rs657152	9	136139265	T	1.7×10^{-8}	*ABO*	0.38	0.023 (0.0062)	1.7×10^{-4}	0.28	0.016 (0.0038)	1.5×10^{-5}

Chr, Chromosome; SE, Standard error. Beta values based on log transformed, adjusted VWFpp values. *Coordinates are in NCBI36/hg18. †In the combined set of 934 individuals in the GABC and 2304 individuals in the TSS cohort.
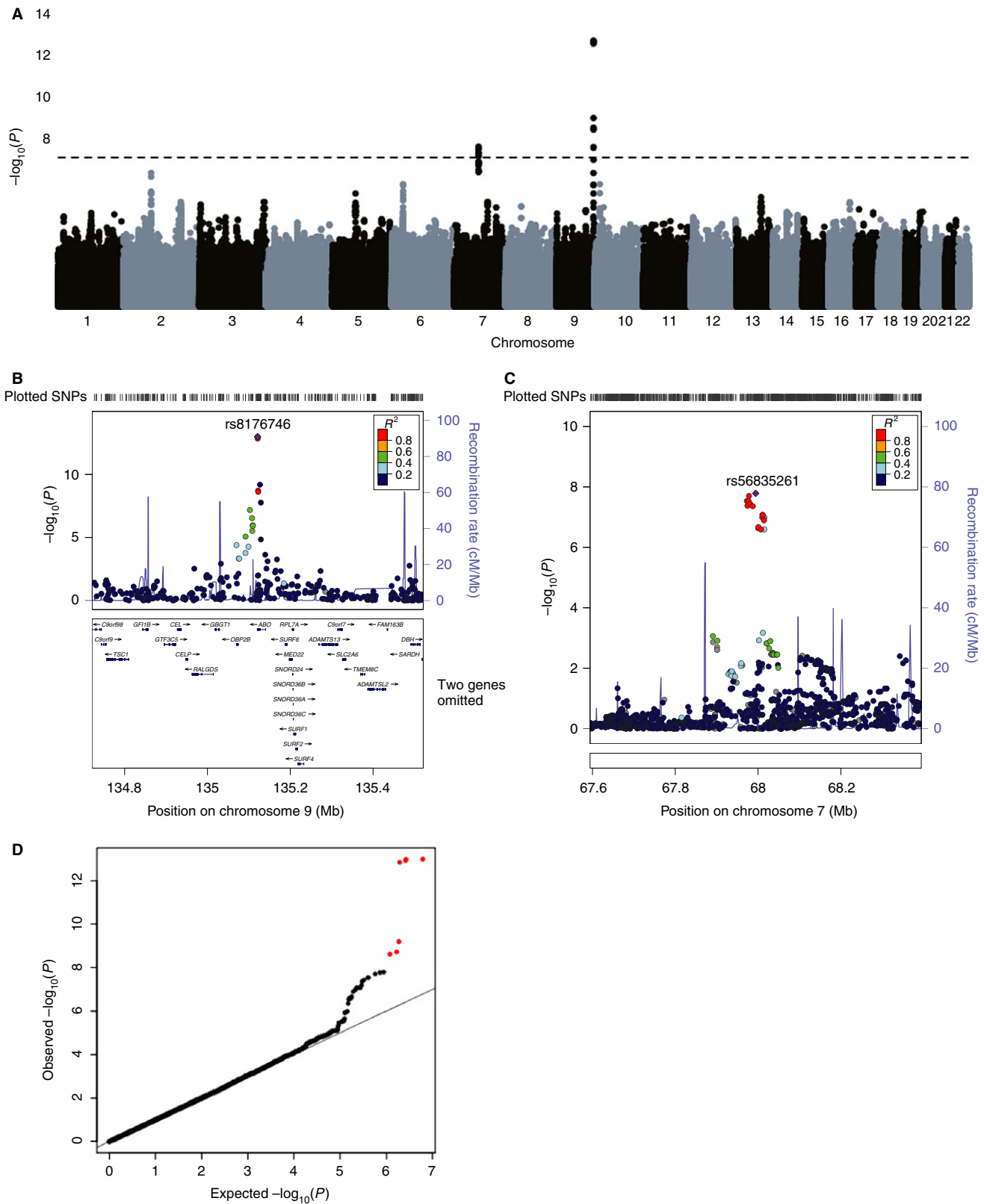
Fig. 1. von Willebrand factor propeptide meta-analysis results. (A) Genome-wide plot of $-\log_{10}(P)$ values for ~ 4.51 million single-nucleotide polymorphisms (SNPs). The dotted line marks the 1.10×10^{-8} threshold of genome-wide significance. (B) Regional plot for the associated region on chromosome 9q34. (C) Regional plot for the associated region on chromosome 7q11.22. (D) Quantile–quantile plot of observed versus expected $-\log_{10}(P)$ values for von Willebrand factor meta-analysis. The observed $-\log_{10}(P)$ values > 7.96 are shown in red.

VWF levels, with Spearman's rank correlations (ρ) of 0.52 for the GABC cohort ($P < 0.0001$) and 0.53 ($P < 0.0001$) for the TSS cohort, respectively (Fig. S5AB).

Common variants at ABO and Chr 7q11 are associated with VWFpp levels

VWFpp association results for the GABC and TSS cohorts (Fig. S3D and Fig. S6A, respectively) were combined in a meta-analysis using a common set of ~ 4.51 million genotyped and imputed SNPs. Two regions with significant associations were identified (Table 2). First, eight SNPs on Chr 9q34 at the *ABO* locus showed significant associations (Fig. 1A,B,D), with the top SNP, rs8176746 ($P = 1.1 \times 10^{-13}$), encoding an L266M substitution in *ABO* and tagging the common *ABO* B allele. All SNPs at the *ABO* locus ($N = 28$) explained 7.40% of the variation in the adjusted VWFpp levels as compared with 19.6% of the variation in the adjusted VWF levels (Table 3). Second, seven associated SNPs were identified on Chr 7q11.22 in an intergenic region, with the top SNP being rs56835261 ($P = 1.6 \times 10^{-8}$) (Fig. 1C). Taken together, the 15 significant SNPs at these two loci explained 3.50% of the variation in the adjusted VWFpp levels (Table 3). A heterogeneity analysis of these top SNPs showed no significant difference across the two cohorts (all with $I^2 = 0\%$ and $P > 0.05$). When a conditional analysis using the top meta-analysis SNP, i.e. rs8176746, was performed, no new signal arose at the *ABO* locus or elsewhere, and the signal at Chr 7q11, i.e. rs56835261, was reduced to just below the threshold for significance ($P = 5.5 \times 10^{-8}$).

ABO haplotypes are associated with VWFpp levels, but less strongly than with VWF levels

Twenty-eight SNPs at the *ABO* locus explained 7.40% of the variation in VWFpp levels as compared with 19.6% of the variation in VWF levels (Table 3) [3]. A1, B, O and A2 tagging haplotypes had the same direction of effect in both VWFpp and VWF associations, but the effect sizes were weaker for VWFpp than for VWF (Table 4). The difference between the sizes of effects of the *ABO* haplotype on VWF levels and VWFpp levels is

evident when the adjusted VWF and VWFpp levels are plotted according to *ABO* haplotypes predicted by three *ABO* SNPs [35] (Fig. 2). There were significant differences in VWF levels between the homozygous low group (A2/O, O/O, and A2/A2) and the homozygous high group (B/B, A1/A1, and B/A1) (t-test, $P = 4.4 \times 10^{-55}$), and the homozygous high group and the heterozygous high/low group (A2/B, B/O, A1/O, and A1/A2) (t-test, $P = 2.8 \times 10^{-10}$) (Fig. 2A). For VWFpp levels, significant differences existed between the homozygous low group and the homozygous high group (t-test, $P = 5.0 \times 10^{-6}$), and the heterozygous high/low group and the homozygous low group (t-test, $P = 5.6 \times 10^{-9}$), but not between the homozygous high group and the heterozygous high/low group (t-test, $P = 0.14$) (Fig. 2B).

No significant linkage signal at Chr 2q12 for VWFpp

Using the sibling subset of GABC and linkage disequilibrium-based 36 658 SNP clusters, we performed linkage studies to identify additional genetic factors affecting VWFpp levels. Unlike VWF, which had a significant linkage signal at Chr 2q12 and Chr 9q34 (containing *ABO*) [3] (Fig. 3A), VWFpp had no significant signal in the Chr 2q12 VWF linkage interval, or at any loci elsewhere (Fig. 3B).

Comparisons of association results of VWF and VWFpp suggest differential clearance mechanisms for the ABO, VWF and Chr 7 regions

If a DNA variant affected the clearance process for either VWF or VWFpp, we would expect it to show different association signals for VWF and VWFpp. Conversely, if a variant affected the shared synthesis or secretion processes for VWF and VWFpp, it would be expected to generate similar association effect sizes and directions for both VWF and VWFpp. Consistent with this expectation, rs687289 and rs8176746, tagging the O and B alleles of *ABO*, respectively, had a strong association with VWF and a much weaker association with VWFpp, suggesting that, although these *ABO* SNPs were associated with both VWF traits, the difference in the magnitude of their effects is best explained by unequal SNP effects on the rates of clearance from the circulation. This result agrees with the well-described ABO blood group association with VWF clearance, but is a new finding for VWFpp (Table 5; Fig. S7A,B).

The top SNP at the *VWF* locus was rs1063856, which encodes a non-synonymous variant (T789P) in the D′ domain of VWF. However, in the VWFpp analysis, no *VWF* SNPs reached genome-wide significance. rs1063856 had the same effect directions for both VWF and VWFpp, but the effect size was greater for VWF (Table 5; Fig. S7C), again consistent with a variant primarily affecting a VWF clearance mechanism.

Table 3 Variance of adjusted VWFpp levels calculated by GCTA and explained by combined GABC and TSS cohorts with imputed SNPs ($n = 3238$)

Region	VWF	VWFpp
Genome-wide (All SNPs)	62.6	56.2
All Significant SNPs*	21.1	3.50
ABO† (28 SNPs)	19.6	7.40
VWF† (415 SNPs)	2.87	1.90

*Results from meta-analysis, VWFpp (15 SNPs), VWF (129 SNPs).
†All SNPs tested in the gene region.

Table 4 ABO haplotype association results for VWF and VWFpp

Allele	Haplotype*	Allele freq.	VWF		VWFpp	
			Beta	*P* value	Beta	*P* value
O	GGC	0.72	−0.100	4.1×10^{-136}	−0.019	5.5×10^{-9}
A1	AGC	0.16	0.110	1.6×10^{-100}	0.010	1.1×10^{-2}
B	AGT	0.072	0.110	1.4×10^{-47}	0.043	1.2×10^{-13}
A2	AAC	0.053	−0.017	0.045	−0.005	4.2×10^{-2}

*Haplotypes based on rs687289, rs8176704 and rs8176749.
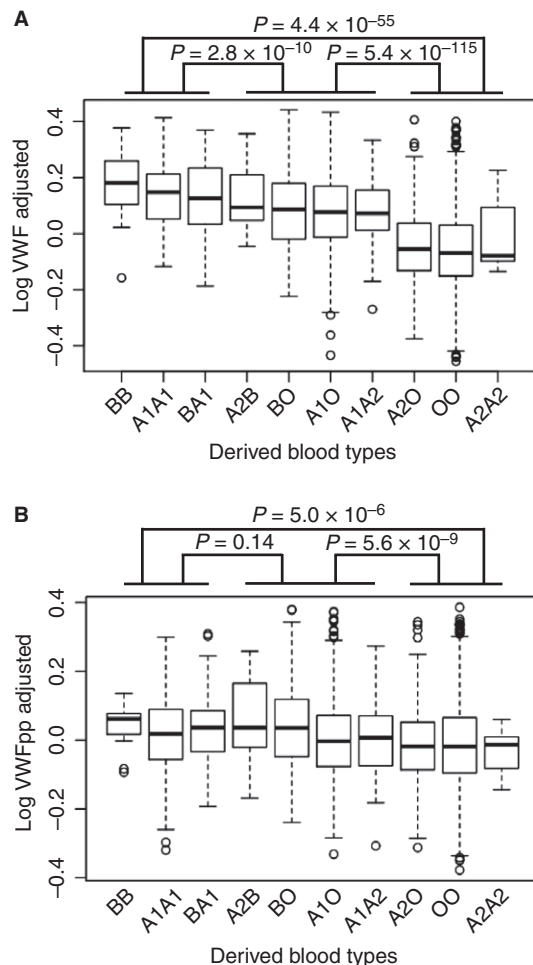
Fig. 2. Distributions of log-transformed and covariate-adjusted von Willebrand factor (VWF) and von Willebrand factor propeptide (VWFpp) levels in the combined Genes and Blood Clotting Study and Trinity Student Study cohorts (N = 3238), grouped by single-nucleotide polymorphism-derived *ABO* haplotype groups. *P*-values show the comparisons between the homozygous high (BB, A1A1, and BA1), heterozygous high/low (A2B, BO, A1O, and A1A2) and homozygous low (A2O, OO, and A2A2) groups (with 'high' and 'low' defined by their adjusted VWF levels).

In the meta-analysis, seven SNPs in the intergenic regions of Chr 7 were associated with VWFpp. These SNPs were not significant in previous VWF GWASs [3,9], and had a lower effect size for VWF, suggesting the presence of a variant affecting the clearance of VWFpp but not that of VWF (Fig. S7D).

Discussion

This study reports the first heritability, linkage and genome-wide association analyses of VWFpp levels, with comparisons with previously published linkage and association studies of plasma VWF measured in the same cohorts. We found that both plasma VWF and VWFpp levels are highly heritable quantitative traits.

In our initial analysis of the VWFpp levels measured with a pair of anti-D2 domain antibodies, we detected a significant association for a non-synonymous SNP in the D2 domain of VWF. However, we determined that this result was probably an artefact resulting from an amino acid substitution affecting antibody binding, and not a genuine association with VWFpp levels. Similar apparent associations with non-synonymous SNPs in cis have been identified for several plasma proteins in the hemostatic system [3,36–38]. In addition, according to the NHGRI GWAS catalog [39] (accessed on 16 December 2015), of the 150 reported GWASs of quantitative protein traits relying on antibody-based assays, 62 (40%) yielded significant associations with SNPs residing at the locus encoding the protein under investigation. Although these SNPs may tag variants altering the level of the corresponding proteins, our findings with rs1800378 and VWFpp suggest that a subset of these signals may be false positives generated by altered antibody binding affinities for the variant protein. This finding is similar to a report of a common *VWF* SNP causing altered VWF:ristocetin cofactor activity levels, but not changing VWF levels [40]. These observations also resemble the caveats reported for gene expression quantitative trait locus (QTL) analysis, where some cis-expression QTLs (eQTLs) may be attributable to variants in cDNA directly affecting hybridization to microarray probes [41,42]. The implication for the genomics community is that extra caution is warranted in the interpretation of cis-QTLs for antibody-based protein traits. We suggest that non-synonymous SNP associations should be replicated with alternative antibody reagents.

Many GWASs have found significant associations with variants in intergenic regions that presumably regulate gene expression. This study identified an association between VWFpp and 14 SNPs in an intergenic region on Chr 7 that were ~ 725 kb upstream from the nearest coding sequence of *AUTS2*. None of these SNPs affect gene expression, according to an eQTL database [43] (accessed
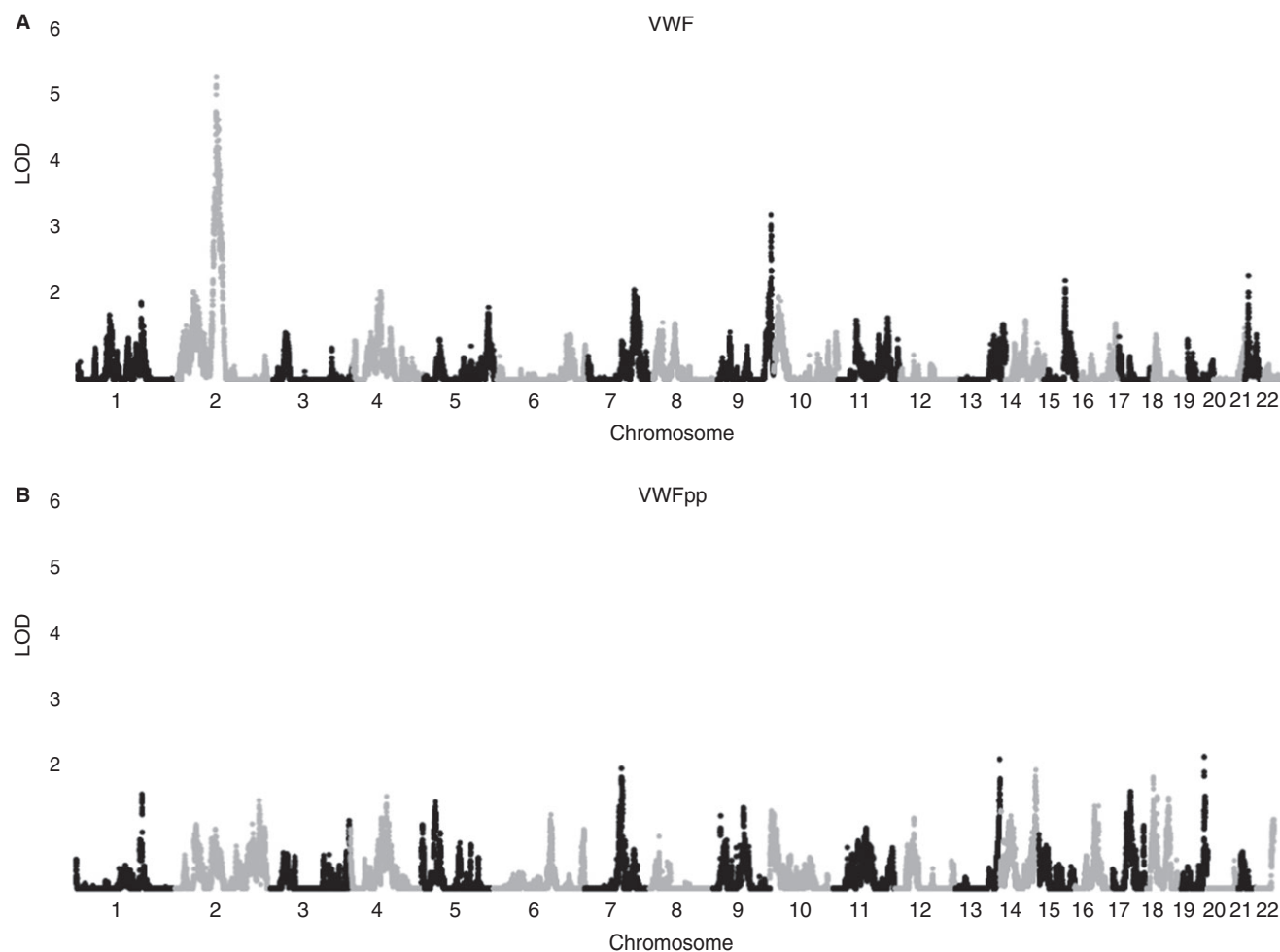
Figure 3. Linkage analysis results in Trinity Student Study and Genes and Blood Clotting Study siblings (*n* = 1065). Genome-wide LOD scores plotted for 36 658 'clusters', defined in MERLIN to model independent regions of linkage. (A) Manhattan plot von Willebrand factor (VWF) linkage results from a previously published analysis [3]. (B) Manhattan plot of von Willebrand factor propeptide (VWFpp) linkage results.

on 7 January 2016). For nine of these SNPs, 'minimal binding evidence' was found by RegulomeDB [44], based on DNase-seq and transcription factor binding motif hits (accessed on 25 January 2016); and 12 of these SNPs had an effect on at least one regulatory motif according to HaploReg [45] (version 4.1, accessed on 25 January 2016). There was no significant association of these SNPs with VWF levels, which suggests that they may regulate the expression of a factor that affects the clearance of VWFpp from the circulation independently of VWF. As these SNPs were detected in the meta-analysis alone, further replication in an independent cohort is necessary for confirmation.

ABO blood group is an established modifier of plasma VWF levels. As other groups have reported [46,47], we found a significant difference in VWF levels between high group homozygotes and high/low group heterozygotes, suggesting an *ABO* allelic dose effect altering steady-state levels of plasma VWF, as opposed to a clean dominant affect of A1-type and B-type antigens, where we would expect to find no differences between homozygous high group and high/low group heterozygotes. Previous studies have failed to detect an association between ABO blood groups and VWFpp levels [12,48–50]. Given the sample size of 3238 individuals, our finding of a significant association was probably attributable to increased power as compared with earlier studies, which had smaller sample sizes, ranging from 47 to 948 individuals. There are four potential N-linked glycosylation sites on VWFpp, but no occupancy of ABO blood group antigens on these sites has been reported [49]. However, Groeneveld *et al.* [11]. recently reported a study of VWF clearance suggesting that the genotype of the individual is the primary determinant of ABO-associated VWF clearance, and not the ABO glycosylation pattern on VWF itself. If this finding applied to VWFpp as well, then the glycosylation pattern on VWFpp would be irrelevant. Nevertheless, the effect of the major *ABO* haplotypes on VWF and VWFpp was in the same direction, suggesting a shared mechanism of action, but functional studies will be required to clarify how ABO antigens alter VWFpp clearance.

Table 5 Comparison of effect size and direction of top meta-analysis single-nucleotide polymorphisms (SNPs) for von Willebrand factor (VWF) and von Willebrand factor propeptide (VWFpp)

SNP (minor allele)	Chr	Locus	VWF β*	VWF *P*-value*	VWFpp β	VWFpp *P*-value
rs687289(A)†	9	*ABO*	0.096	7.7E-138	0.019	1.3E-9
rs8176746(T)‡	9	*ABO*	0.11	1.2E-49	0.044	1.6E-13
rs1063856(C)§	12	*VWF*	0.031	6.7E-17	0.014	7.7E-6
rs56835261(A)	7	Intergenic	0.015	3.5E-2	0.035	8.5E-9

SNPs were selected for further analysis if they were the top SNPs defining a locus in the meta-analysis for VWF [3] or VWFpp levels. *Results from previous VWF analysis [3]. †This SNP tags the common O allele of *ABO*. ‡This SNP tags the common B allele of *ABO*. §This SNP encodes a non-synonymous *VWF* variant T789P in the D′ domain.

For the *VWF* SNPs, the difference in effect size between VWF and VWFpp and the absence of significant associations with VWFpp levels suggest that the VWF variants operate through an altered clearance pathway(s). Investigators have cloned and functionally characterized many *VWF* mutations causing VWD [51], but common *VWF* variants causing variations in VWF levels in healthy individuals have not yet been identified or well characterized. Previous studies by our group and others have documented an association of VWF levels with common variants at the *VWF* locus [3,9,10]. The top *VWF* SNP in our meta-analysis, i.e. rs1063856(C), encodes a missense variant in the D′ domain of VWF, and could be a functional variant driving the association. Our study suggests that the *VWF* haplotype tagged by this SNP encodes a form of VWF with a prolonged plasma half-life as compared with the VWF encoded by the reference allele, and that the effect of this haplotype on VWFpp clearance is much weaker. The precise mechanism for the altered clearance of this VWF variant remains unknown.

Linkage analysis of VWFpp did not detect a QTL at Chr 2q12, which was previously identified as a QTL for VWF levels. The linkage interval contains many potential candidate genes, and our results suggest that genes potentially affecting protein clearance, such as those encoding sialotransferases (*ST3GAL5*) or lectin receptors (*LMAN2L*), are more attractive candidates than genes likely to affect synthesis or secretion pathways, such as those encoding SNARE complex (*VAMP5* and *VAMP8*) or Golgi-associated (*TGOLN2*) proteins.

Our study used the comparative analysis of VWF and VWFpp to distinguish genetic variants affecting VWF or VWFpp clearance from those affecting synthesis and secretion. However, as a novel strategy to interpret differential associations of two related proteins, our analyses had several limitations. First, by necessity, the measurement of VWFpp was performed with a different pair of antibodies from those employed in the VWF assay. This

allows for differences in assay performance that may have led to false-negative associations of VWFpp, suggesting differential clearance mechanisms. However, this scenario is not very likely, as the heritability of VWFpp levels was very similar to that of VWF levels, and both traits were measured in the same cohorts. Second, variants that were associated with both clearance and synthesis/secretion rates of VWFpp may have been undetectable in our analysis, e.g. a variant associated with both decreased secretion rates and decreased clearance rates. Although our results strongly suggest a clearance mechanism for the major common modifiers of VWF, functional studies will be required for confirmation.

Taken together, the results of our VWFpp analyses provide new insights into the mechanism of action for common variants (at *ABO* and *VWF*) and potentially rare variants (in the Chr 2 lineage region) altering plasma VWF levels, and demonstrate a newly discovered association signal for VWFpp at Chr 7. This study is expected to facilitate the identification of functional variants controlling VWF variation, and improve our understanding of the molecular mechanisms in individuals with bleeding disorders resulting from low VWF levels, or in individuals at risk for venous thromboembolic disease resulting from elevated VWF levels [52].

Discosure of Conflict of Interests

The authors state that they have no conflict of interest.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Epitope mapping anti-VWFpp antibodies.
Fig. S2. Single-donor VWFpp AlphaLISA assay.
Fig. S3. VWFpp association analysis results in GABC.
Fig. S4. Raw and adjusted VWFpp distributions in GABC (934 European individuals) and TSS (2304 individuals) cohorts.
Fig. S5. Scatter plot of individual log-adjusted VWFpp and VWF levels of the (A) GABC and (B) TSS participants.
Fig. S6. VWFpp association analysis results in TSS.
Fig. S7. Distribution of log-transformed VWF and VWFpp by top meta-analysis SNP allelic series.

References

1 Bladbjerg EM, de Maat MP, Christensen K, Bathum L, Jespersen J, Hjelmborg J. Genetic influence on thrombotic risk markers in the elderly – a Danish twin study. *J Thromb Haemost* 2006; **4**: 599–607.

2 de Lange M, Snieder H, Ariens RA, Spector TD, Grant PJ. The genetics of haemostasis: a twin study. *Lancet* 2001; **357**: 101–5.

3 Desch KC, Ozel AB, Siemieniak D, Kalish Y, Shavit JA, Thornburg CD, Sharathkumar AA, McHugh CP, Laurie CC, Crenshaw A, Mirel DB, Kim Y, Cropp CD, Molloy AM, Kirke PN, Bailey-Wilson JE, Wilson AF, Mills JL, Scott JM, Brody LC, *et al.* Linkage analysis identifies a locus for plasma von Willebrand factor undetected by genome-wide association. *Proc Natl Acad Sci USA* 2013; **110**: 588–93

4 Orstavik KH, Magnus P, Reisner H, Berg K, Graham JB, Nance W. Factor VIII and factor IX in a twin population. Evidence for a major effect of ABO locus on factor VIII level. *Am J Hum Genet* 1985; **37**: 89–101.

5 Souto JC, Almasy L, Soria JM, Buil A, Stone W, Lathrop M, Blangero J, Fontcuberta J. Genome-wide linkage analysis of von Willebrand factor plasma levels: results from the GAIT project. *Thromb Haemost* 2003; **89**: 468–74.

6 Sadler JE. von Willebrand factor: two sides of a coin. *J Thromb Haemost*. 2005;**3**:1702–9.

7 Sonneveld MA, de Maat MP, Leebeek FW. Von Willebrand factor and ADAMTS13 in arterial thrombosis: a systematic review and meta-analysis. *Blood Rev* 2014; **28**: 167–78.

8 Sanders YV, van der Bom JG, Isaacs A, Cnossen MH, de Maat MP, Laros-van Gorkom BA, Fijnvandraat K, Meijer K, van Duijn CM, Mauser-Bunschoten EP, Eikenboom J, Leebeek FW, Wi NSG. CLEC4M and STXBP5 gene variations contribute to von Willebrand factor level variation in von Willebrand disease. *J Thromb Haemost* 2015; **13**: 956–66.

9 Smith NL, Chen MH, Dehghan A, Strachan DP, Basu S, Soranzo N, Hayward C, Rudan I, Sabater-Lleal M, Bis JC, de Maat MP, Rumley A, Kong X, Yang Q, Williams FM, Vitart V, Campbell H, Malarstig A, Wiggins KL, Van Duijn CM, *et al.* Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation* 2010; **121**: 1382–92.

10 Campos M, Sun W, Yu F, Barbalic M, Tang W, Chambless LE, Wu KK, Ballantyne C, Folsom AR, Boerwinkle E, Dong JF. Genetic determinants of plasma von Willebrand factor antigen levels: a target gene SNP and haplotype analysis of ARIC cohort. *Blood* 2011; **117**: 5224–30.

11 Groeneveld DJ, van Bekkum T, Cheung KL, Dirven RJ, Castaman G, Reitsma PH, van Vlijmen B, Eikenboom J. No evidence for a direct effect of von Willebrand factor's ABH blood group antigens on von Willebrand factor clearance. *J Thromb Haemost* 2015; **13**: 592–600.

12 Gallinaro L, Cattini MG, Sztukowska M, Padrini R, Sartorello F, Pontara E, Bertomoro A, Daidone V, Pagnan A, Casonato A. A shorter von Willebrand factor survival in O blood group subjects explains how ABO determinants influence plasma von Willebrand factor. *Blood* 2008; **111**: 3540–5.

13 O'Donnell J, Boulton FE, Manning RA, Laffan MA. Amount of H antigen expressed on circulating von Willebrand factor is modified by ABO blood group genotype and is a major determinant of plasma von Willebrand factor antigen levels. *Arterioscler Thromb Vasc Biol* 2002; **1**: 335–41.

14 Mohlke KL, Purkayastha AA, Westrick RJ, Smith PL, Petryniak B, Lowe JB, Ginsburg D. Mvwf, a dominant modifier of murine von Willebrand factor, results from altered lineage-specific expression of a glycosyltransferase. *Cell* 1999; **96**: 111–20.

15 Springer TA. Biology and physics of von Willebrand factor concatamers. *J Thromb Haemost* 2011; **9**(Suppl. 1): 130–43.

16 Haberichter SL, Balistreri M, Christopherson P, Morateck P, Gavazova S, Bellissimo DB, Manco-Johnson MJ, Gill JC, Montgomery RR. Assay of the von Willebrand factor (VWF) propeptide to identify patients with type 1 von Willebrand disease with decreased VWF survival. *Blood* 2006; **108**: 3344–51.

17 Haberichter SL, Castaman G, Budde U, Peake I, Goodeve A, Rodeghiero F, Federici AB, Batlle J, Meyer D, Mazurier C, Goudemand J, Eikenboom J, Schneppenheim R, Ingerslev J, Vorlova Z, Habart D, Holmberg L, Lethagen S, Pasi J, Hill FG, *et al.* Identification of type 1 von Willebrand disease patients with reduced von Willebrand factor survival by assay of the VWF propeptide in the European study: molecular and clinical markers for the diagnosis and management of type 1 VWD (MCMDM-1VWD). *Blood* 2008; **111**: 4979–85.

18 Haberichter SL. von Willebrand factor propeptide: biology and clinical utility. *Blood* 2015; **8**: 1753–61.

19 Ma Q, Ozel AB, Ramdas S, McGee B, Khoriaty R, Siemieniak D, Li HD, Guan Y, Brody LC, Mills JL, Molloy AM, Ginsburg D, Li JZ, Desch KC. Genetic variants in PLG, LPA, and SIGLEC 14 as well as smoking contribute to plasma plasminogen levels. *Blood* 2014; **124**: 3155–64.

20 Desch K, Li J, Kim S, Laventhal N, Metzger K, Siemieniak D, Ginsburg D. Analysis of informed consent document utilization in a minimal-risk genetic study. *Ann Intern Med* 2011; **155**: 316–22.

21 Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, Hao L, Kiang A, Paschall J, Phan L, Popova N, Pretel S, Ziyabari L, Lee M, Shao Y, Wang ZY, Sirotkin K, Ward M, Kholodov M, Zbicz K, *et al.* The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet* 2007; **39**: 1181–6.

22 Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet* 2009; **84**: 210–23.

23 Delaneau O, Howie B, Cox AJ, Zagury JF, Marchini J. Haplotype estimation using sequencing reads. *Am J Hum Genet* 2013; **93**: 687–96.

24 Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012; **44**: 955–9.

25 Zhu X, Li S, Cooper RS, Elston RC. A unified association analysis approach for family and unrelated samples correcting for stratification. *Am J Hum Genet* 2008; **82**: 352–65.

26 Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. Variance component model to account

for sample structure in genome-wide association studies. *Nat Genet* 2010; **42**: 348–54.

27 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–75.

28 Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999; **55**: 997–1004.

29 Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; **26**: 2190–1.

30 Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010; **26**: 2336–7.

31 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010; **7**: 248–9.

32 Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin – rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30**: 97–101.

33 Ihaka R, Gentleman R. R: a language for data analysis and graphics. *J Comp Graph Stat* 1996; **5**: 299–314.

34 Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**: 76–82.

35 Barbalic M, Dupuis J, Dehghan A, Bis JC, Hoogeveen RC, Schnabel RB, Nambi V, Bretler M, Smith NL, Peters A, Lu C, Tracy RP, Aleksic N, Heeriga J, Keaney Jr JF, Rice K, Lip GY, Vasan RS, Glazer NL, Larson MG, *et al.* Large-scale genomic studies reveal central role of ABO in sP-selectin and sICAM-1 levels. *Hum Mol Genet* 2010; **19**: 1863–72.

36 de Vries PS, Boender J, Sonneveld MA, Rivadeneira F, Ikram MA, Rottensteiner H, Hofman A, Uitterlinden AG, Leebeek FW, Franco OH, Dehghan A, de Maat MP. Genetic variants in the ADAMTS13 and SUPT3H genes are associated with ADAMTS13 activity. *Blood* 2015; **125**: 3949–55.

37 Huang J, Sabater-Lleal M, Asselbergs FW, Tregouet D, Shin S-Y, Ding J, Baumert J, Oudot-Mellakh T, Folkersen L, Johnson AD, Smith NL, Williams SM, Ikram MA, Kleber ME, Becker DM, Truong V, Mychaleckyj JC, Tang W, Yang Q, Sennblad B, *et al.* Genome-wide association study for circulating levels of PAI-1 provides novel insights into its regulation. *Blood* 2012; **2012**: 4873–81.

38 Tang W, Basu S, Kong X, Pankow JS, Aleksic N, Tan A, Cushman M, Boerwinkle E, Folsom AR. Genome-wide association study identifies novel loci for plasma levels of protein C: the ARIC study. *Blood* 2010; **116**: 5032–6.

39 Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L, Parkinson H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 2014; **42**: D1001–6.

40 Flood VH, Gill JC, Morateck PA, Christopherson PA, Friedman KD, Haberichter SL, Branchford BR, Hoffmann RG, Abshire TC, Di Paola JA, Hoots WK, Leissinger C, Lusher JM, Ragni MV, Shapiro AD, Montgomery RR. Common VWF exon 28 polymorphisms in African Americans affecting the VWF activity assay by ristocetin cofactor. *Blood* 2010; **116**: 280–6.

41 Sliwerska E, Meng F, Speed TP, Jones EG, Bunney WE, Akil H, Watson SJ, Burmeister M. SNPs on chips: the hidden genetic code in expression arrays. *Biol Psychiatry* 2007; **1**: 13–16.

42 Benovoy D, Kwan T, Majewski J. Effect of polymorphisms within probe-target sequences on olignonucleotide microarray experiments. *Nucleic Acids Res* 2008; **36**: 4417–23.

43 Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M, Pritchard JK. High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet* 2008; **4**: e1000214.

44 Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; **22**: 1790–7.

45 Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012; **40**: D930–4.

46 Souto JC, Almasy L, Muniz-Diaz E, Soria JM, Borrell M, Bayen L, Mateo J, Madoz P, Stone W, Blangero J, Fontcuberta J. Functional effects of the ABO locus polymorphism on plasma levels of von Willebrand factor, factor VIII, and activated partial thromboplastin time. *Arterioscler Thromb Vasc Biol* 2000; **20**: 2024–8.

47 Shima M, Fujimura Y, Nishiyama T, Tsujiuchi T, Narita N, Matsui T, Titani K, Katayama M, Yamamoto F, Yoshioka A. ABO blood group genotype and plasma von Willebrand factor in normal individuals. *Vox Sang* 1995; **68**: 236–40.

48 Eikenboom J, Federici AB, Dirven RJ, Castaman G, Rodeghiero F, Budde U, Schneppenheim R, Batlle J, Canciani MT, Goudemand J, Peake I, Goodeve A. VWF propeptide and ratios between VWF, VWF propeptide, and FVIII in the characterization of type 1 von Willebrand disease. *Blood* 2013; **121**: 2336–9.

49 Nossent AY, Van Marion V, Van Tilburg NH, Rosendaal FR, Bertina RM, Van Mourik JA, Eikenboom HC. von Willebrand factor and its propeptide: the influence of secretion and clearance on protein levels and the risk of venous thrombosis. *J Thromb Haemost* 2006;**4**:2556–62.

50 Marianor M, Zaidah AW, Maraina ChC. von Willebrand factor propeptide: a potential disease biomarker not affected by ABO blood groups. *Biomarker Insights* 2015; **10**: 75–9.

51 Hampshire DJ, Goodeve AC. The International Society on Thrombosis and Haematosis von Willebrand disease database: an update. *Semin Thromb Hemost* 2011; **37**: 470–9.

52 Smith NL, Rice KM, Bovill EG, Cushman M, Bis JC, McKnight B, Lumley T, Glazer NL, van Hylckama Vlieg A, Tang W, Dehghan A, Strachan DP, O'Donnell CJ, Rotter JI, Heckbert SR, Psaty BM, Rosendaal FR. Genetic variation associated with plasma von Willebrand factor levels and the risk of incident venous thrombosis. *Blood* 2011; **117**: 6007–11.