

## Article

# Blood Levels of the *SMOC1* Hepatokine Are Not Causally Linked with Type 2 Diabetes: A Bidirectional Mendelian Randomization Study

Nooshin Ghodsian <sup>1</sup>, Eloi Gagnon <sup>1</sup>, Jérôme Bourgault <sup>1</sup>, Émilie Gobeil <sup>1</sup> , Hasanga D. Manikpurage <sup>1</sup> ,  
Nicolas Perrot <sup>1</sup>, Arnaud Girard <sup>1</sup>, Patricia L. Mitchell <sup>1</sup>  and Benoit J. Arsenault <sup>1,2,\*</sup> 

<sup>1</sup> Centre de Recherche de l'Institut Universitaire de Cardiologie et de Pneumologie de Québec, Québec, QC G1V 4G5, Canada; nooshin.ghodsian.1@ulaval.ca (N.G.); eloi.gagnon.1@ulaval.ca (E.G.); jerome.bourgault@criucpq.ulaval.ca (J.B.); emilie.gobeil.2@ulaval.ca (É.G.); hasanga.manik-purage.1@ulaval.ca (H.D.M.); nicolas.perrot@criucpq.ulaval.ca (N.P.); arnaud.girard.1@ulaval.ca (A.G.); Patricia.Mitchell@criucpq.ulaval.ca (P.L.M.)

<sup>2</sup> Department of Medicine, Faculty of Medicine, Université Laval, Québec, QC G1V 0A6, Canada

\* Correspondence: benoit.arsenault@criucpq.ulaval.ca; Tel.: +1-418-656-8711 (ext. 3498)

**Abstract:** Hepatokines are liver-derived proteins that may influence metabolic pathways such as insulin sensitivity. Recently, Sparc-related modular calcium-binding protein 1 (*SMOC1*) was identified as glucose-responsive hepatokine that is dysregulated in the setting of non-alcoholic fatty liver disease (NAFLD). While *SMOC1* may influence glucose-insulin homeostasis in rodents, it is unknown if *SMOC1* is influenced by NAFLD in humans. It is also unknown if *SMOC1* is causally associated with metabolic and disease traits in humans. Therefore, we aimed to determine the effect of NAFLD on *SMOC1* gene expression in the liver and aimed to explore the potential causal associations of *SMOC1* levels with NAFLD, T2D, and glycemic traits in humans. Using an RNA sequencing dataset from a cohort of 216 patients with NAFLD, we assessed *SMOC1* expression levels across the NAFLD spectrum. We performed a series of bidirectional inverse-variance weighted Mendelian randomization (MR) analyses on blood *SMOC1* levels using two sources of genome-wide association studies (GWAS) (Fenland study, n = 10,708 and INTERVAL study, n = 3301). We utilized GWAS summary statistics for NAFLD in 8434 cases and 770,180 controls, as well as publicly available GWAS for type 2 diabetes (T2D), body mass index (BMI), waist-to-hip ratio (WHR), fasting blood insulin (FBI), fasting blood glucose (FBG), homeostatic Model Assessment of Insulin Resistance (HOMA-B and HOMA-IR), and hemoglobin A1c (HbA1C). We found that *SMOC1* expression showed no significant differences across NAFLD stages. We also identified that the top single-nucleotide polymorphism associated with blood *SMOC1* levels, was associated with *SMOC1* gene expression in the liver, but not in other tissues. Using MR, we did not find any evidence that genetically predicted NAFLD, T2D, and glycemic traits influenced *SMOC1* levels. We also did not find evidence that blood *SMOC1* levels were causally associated with T2D, NAFLD, and glycemic traits. In conclusion, the hepatokine *SMOC1* does not appear to be modulated by the presence of NAFLD and may not regulate glucose-insulin homeostasis in humans. Results of this study suggest that blood factors regulating metabolism in rodents may not always translate to human biology.

**Keywords:** *SMOC1*; type 2 diabetes; Mendelian randomization; hepatokine



**Citation:** Ghodsian, N.; Gagnon, E.; Bourgault, J.; Gobeil, É.; Manikpurage, H.D.; Perrot, N.; Girard, A.; Mitchell, P.L.; Arsenault, B.J. Blood Levels of the *SMOC1* Hepatokine Are Not Causally Linked with Type 2 Diabetes: A Bidirectional Mendelian Randomization Study. *Nutrients* **2021**, *13*, 4208. <https://doi.org/10.3390/nu13124208>

Academic Editor: Omorogieva Ojo

Received: 25 October 2021

Accepted: 20 November 2021

Published: 24 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Hepatokines are liver-secreted proteins that influence a wide range of biological processes such as lipoprotein-lipid metabolism, glucose-insulin homeostasis, inflammation, coagulation, fibrinolysis, etc. [1,2]. Therefore, hepatokines may be involved in several diseases such as non-alcoholic fatty liver disease (NAFLD), type 2 diabetes (T2D), and atherosclerotic cardiovascular diseases [3,4]. Hepatic steatosis may also induce changes in

hepatokine secretion patterns and the promotion of insulin resistance [2]. A recent study indicated that the mouse liver could be responsible for the secretion of up to 538 proteins, of which 71 were dysregulated in the setting of NAFLD [5]. Sparc-related modular calcium-binding protein 1 (SMOC1) was identified as one of those hepatokines involved in glucose-insulin metabolism in primary murine hepatocytes [4]. *SMOC1* is highly expressed and secreted by the liver and regulates glucose homeostasis [6].

In mice, *SMOC1* infusions and adenoviral-mediated overexpression of *SMOC1* in the liver, significantly improved glycemic control and histological features of NAFLD, suggesting that *SMOC1* might be a potential therapeutic agent for the treatment of T2D [4,7]. Although lower *SMOC1* levels in the blood have been demonstrated in the setting of insulin resistance and obesity in humans, the causal role of *SMOC1* on the etiology of T2D and obesity in humans has not been investigated. It is therefore unknown if *SMOC1* represents a therapeutic target for the management of T2D in humans. It also remains to be determined if NAFLD in humans is associated with altered liver gene expression of *SMOC1* or alterations in the concentrations of *SMOC1* in the blood.

Mendelian randomization (MR) is a causal inference method relying on the random allocation of alleles at conception to estimate causal effects on outcomes. MR is an increasingly recognized method to determine whether genetically regulated exposures (such as blood factors) are causally linked with outcomes (such as human diseases and metabolic traits). MR has been used to determine the association between hepatokine levels and human diseases and traits. For instance, Thakkinstian et al. showed a causal association between circulating fetuin-A and BMI [8]. Other studies revealed associations between serum testosterone, sex hormone-binding globulin and several disease traits such as T2D [9–11].

Here, we aimed to determine whether the effect of *SMOC1* on glycemic traits in mice could be extended to humans. First, we investigated whether liver *SMOC1* expression was altered in the setting of human NAFLD. Second, using bidirectional MR, we sought to determine whether blood levels of *SMOC1* were causally influenced by NAFLD, obesity or glycemic traits and if *SMOC1* could be causally implicated in the etiology of metabolic traits and metabolic diseases such as NAFLD and T2D.

## 2. Methods

### 2.1. Observational Analysis

We first evaluated gene expression levels of *SMOC1* across different stages of NAFLD using *SMOC1* RNA expression from a cohort of 216 patients (206 NAFLD cases and 10 healthy controls) across the full histologic range of NAFLD [12]. In order to provide an accurate estimation of gene expression, we normalized the raw data from RNA sequencing with the trimmed mean of M values (TMM) method [13] and obtained gene weights using limma's voom methodology [14]. We then performed one-way ANOVA tests and Tukey multiple comparisons of means (95% family-wise confidence level) to evaluate differences in *SMOC1* gene expression levels between two or more NAFLD groups. We compared control, NAFL, and the spectrum of NASH classifications including NASH-F0-F1 (no fibrosis), NASH-F2 (moderate fibrosis), NASH-F3 (advanced fibrosis) and NASH-F4 (severe fibrosis/cirrhosis).

### 2.2. Study Populations Included in the Mendelian Randomization Analyses

We obtained information from publicly accessible GWAS summary statistics of European ancestry with no sample overlap between exposures and outcomes. *SMOC1*: We extracted the effect of single-nucleotide polymorphisms (SNPs) in two GWAS of blood *SMOC1* levels measured in 3301 individuals of the INTERVAL study [15], and 10,708 individuals from the Fenland cohort [16]. In the INTERVAL study, Sun et al. performed genome-wide genotyping of 10.6 million imputed autosomal variants against levels of 2994 plasma proteins in 3301 individuals of European descent. The relative concentrations of 3622 plasma proteins or protein complexes were assayed using 4034 modified aptamers (SomaSCAN). In the Fenland cohort, Pietzner et al. also integrated large-scale

genomic and aptamer-based plasma proteomic data from a population-based study of 10,708 individuals to characterize the genetic architecture of 179 host proteins. They characterized protein quantitative trait loci (pQTLs) in close proximity to protein-encoding genes ( $\pm 500$  Kb window around the gene body), cis-pQTLs. NAFLD: GWAS summary statistics for NAFLD were obtained from four cohorts including the Electronic Medical Records and Genomics (eMERGE) network, the UK Biobank, Estonian Biobank, and FinnGen with 8434 NAFLD cases and 770,180 controls [17]. T2D: Data were extracted from combined genome-wide association data of 32 studies, including 74,124 T2D cases and 824,006 controls of European ancestry [18]. Body mass index (BMI) and waist-to-hip ratio (WHR): Summary statistics were obtained from the Genetic Investigation of Anthropometric Traits (GIANT) consortium. Summary statistics for BMI were obtained from a meta-analysis of up to 125 GWAS for 339,224 European individuals [19]. Summary statistics for WHR were obtained from a meta-analysis of 210,088 individuals [20]. Glycemic traits: Summary statistics were obtained from the Meta-Analyses of Glucose and Insulin-related traits consortium (MAGIC). Summary statistics for fasting insulin (FI) and fasting glucose (FG) were obtained in 51,750 and 58,074 non-diabetic individuals respectively [21]. Summary statistics for Homeostatic Model Assessment of Insulin Resistance (HOMA-B and HOMA-IR) were obtained from 36,466 for HOMA-B and 37,037 for HOMA-IR individuals [22]. Summary statistics for hemoglobin A1c (HbA1C) were obtained from 46,368 individuals [23]. Units for continuous traits are in standard deviation (SD) and for dichotomous traits such as disease status in log (OR).

### 2.3. Mendelian Randomization Analyses

We applied MR to assess the causal effect of NAFLD, T2D, obesity and glycemic traits on blood *SMOC1* levels. We selected independent ( $r^2 < 0.001$ ) genome-wide significant SNPs ( $p < 5 \times 10^{-8}$ ) associated with these traits. We used a  $p$ -value threshold of  $5 \times 10^{-6}$  for HOMA-IR because there was no genome-wide significant instrument. We performed inverse variance weighted (IVW) analysis as primary MR analysis. For an association that reached nominal significance, we performed 6 different robust MR methods: the MR-Robust Adjusted Profile Score (MR-RAPS) [24], the contamination mixture [25], the weighted median, the weighted mode, the MR-Egger [26] and the MR-PRESSO [27] approaches. Consistent estimates across these methods provide further confirmation about the nature of the causal links. To quantify instrument strength, we used the F-statistic [28]. To quantify the variance explained we used the  $R^2$  [29] and to quantify estimates heterogeneity we used Cochran's Q statistic [30].

We applied three types of MR analysis to assess the causal effect of liver *SMOC1* expression on glycemic traits. First, we used cis-MR analysis with the sentinel SNP (the one with the strongest effect on *SMOC1* levels) to calculate the Wald ratio. We selected independent ( $r^2 < 0.001$ ) genome-wide significant SNPs ( $p < 5 \times 10^{-8}$ ) in cis, that is 500 Kb downstream and 500 Kb upstream, of the *SMOC1* gene and performed MR analysis. Cis-acting pQTLs (close to *SMOC1*) are more specific instruments, suppressing or upregulating the expression of a gene, whereas trans pQTLs (distal to the gene) may operate via more complex mechanisms and are therefore more likely to be pleiotropic. Second, we used pan (cis + trans) MR analysis. We selected independent ( $r^2 < 0.001$ ) genome-wide significant SNPs ( $p < 5 \times 10^{-8}$ ) from all regions of the genome and performed MR analysis. The inclusion of trans-acting pQTL increased the number of genetic instruments, thereby increasing the power to detect association. This also allowed the use of robust analyses and the quantification of heterogeneity (Cochran's Q) to assess the validity of the MR assumptions. The inclusion of trans pQTL, however, can introduce pleiotropy. Third, we performed multi-cis MR analyses. We included as genetic instruments all cis-acting SNP moderately independently (LD clumping  $r^2 < 0.6$ ) associated at  $p < 5 \times 10^{-8}$ . We used a generalized inverse-variance weighted (IVW) model that take into account LD between SNPs. LD correlation matrix were obtained from the 1000 Genomes European ancestry reference samples [31]. This method can result in more precise estimates [32]. Altogether,

the use of different MR analyses with different strengths and weaknesses increases the robustness of the causal finding.

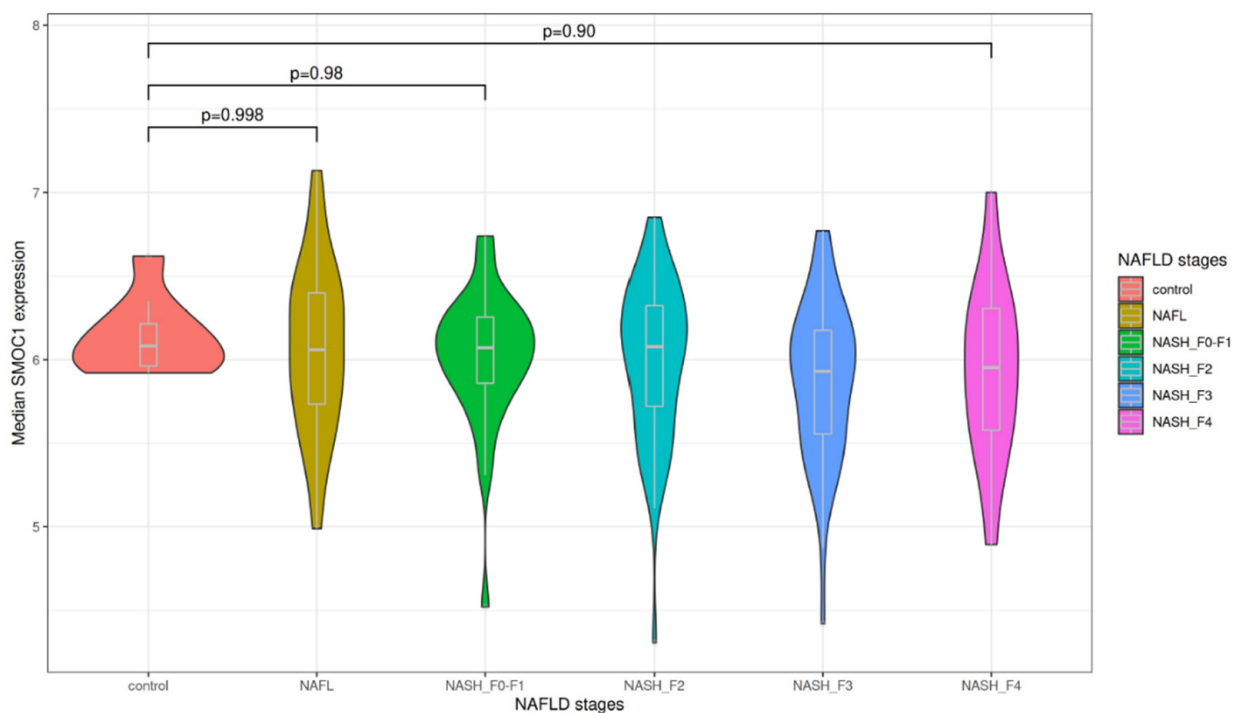
#### 2.4. Data and Code Availability

All GWAS summary statistics used in this paper are publicly available. The code used to generate these results is uploaded on GitHub: <https://github.com/LaboArsenault> (accessed on 25 august 2021).

### 3. Results

#### 3.1. Association of Liver *SMOC1* Expression with Liver Disease Progression

Differences in *SMOC1* gene expression levels across stages of steatohepatitis (control, NAFL, NASH-F0-F1, NASH-F2, NASH-F3 and NASH-F4) are presented in Figure 1. One-way ANOVA revealed that there was no significant difference in *SMOC1* RNA expression between the groups. This observation was further confirmed by Tukey's Honestly Significant Difference (Tukey's HSD) post-hoc test for pairwise comparisons. Liver *SMOC1* gene expression levels are therefore not associated with the progression of fatty liver disease.

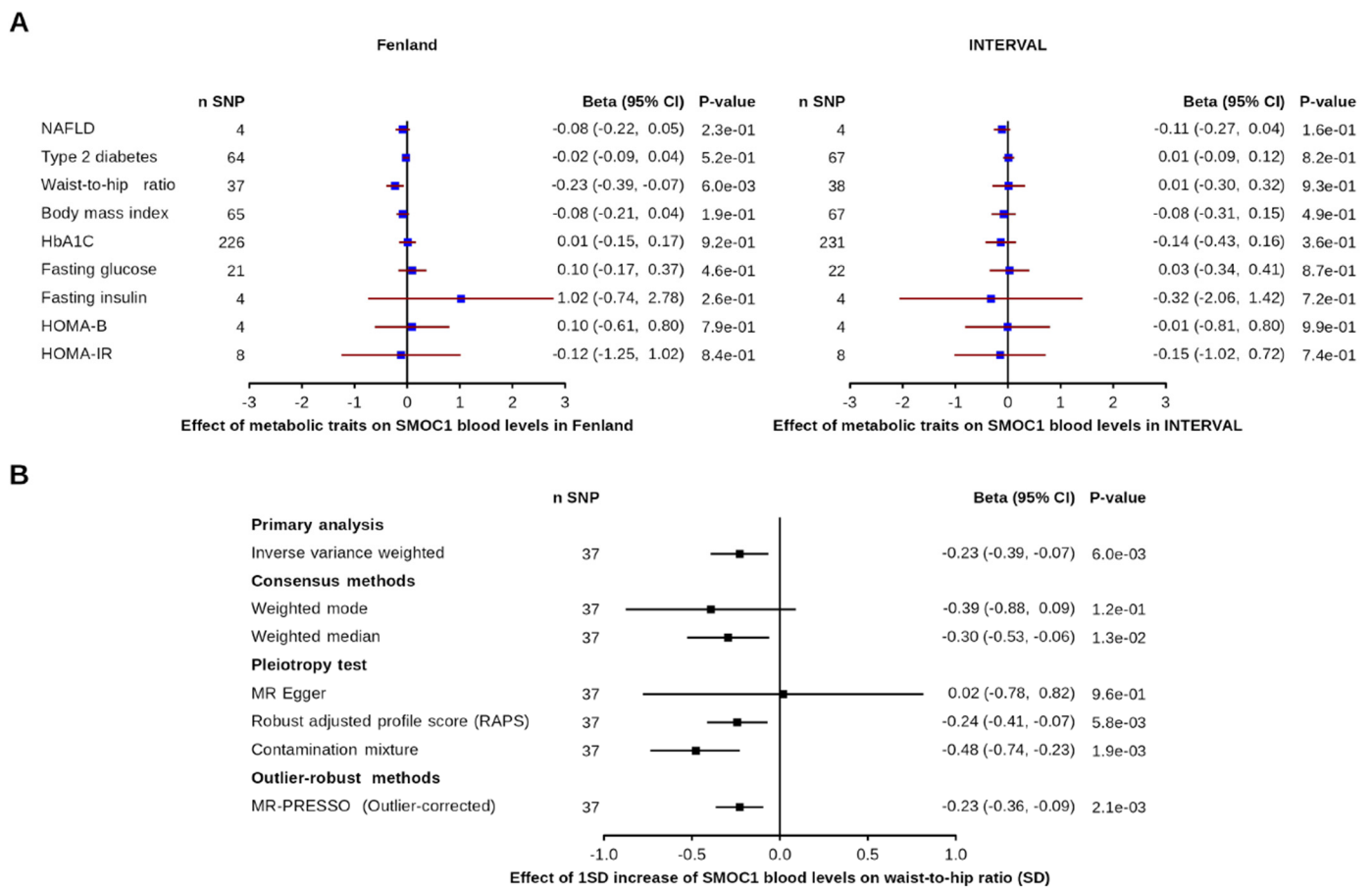


**Figure 1.** *SMOC1* gene expression levels across non-alcoholic fatty liver disease (NAFLD) stages. The violin plot indicates the mean level of *SMOC1* RNA expression across stages of steatohepatitis. This data from one-way ANOVA shows no significant difference of *SMOC1* RNA expression between any groups (control, NAFL, NASH-F0-F1, NASH-F2, NASH-F3, and NASH-F4).

#### 3.2. Effect of Metabolic and Disease-Related Traits on *SMOC1* Levels

The effect of metabolic diseases and glycemic traits on *SMOC1* are presented in Figure 2. Only the association of WHR with *SMOC1* in the Fenland dataset reached statistical significance, although the same association in the INTERVAL dataset did not (Figure 2 panel A). Specifically, each standard deviation increases in WHR decreased *SMOC1* blood levels by  $-0.23$  (95% CI =  $-0.39, -0.066$ ,  $p$ -value =  $6 \times 10^{-3}$ ) standard deviation. Cochran's Q test indicated no significant heterogeneity across estimates ( $Q = 24.67$ ,  $p$ -value =  $0.92$ ). Robust MR analysis indicated low pleiotropy as most (4/6) robust MR analyses reached nominal significance (Figure 2 panel B). These results support that waist circumference might be linked with lower *SMOC1* levels, although this result did not replicate in another

dataset. Altogether, we found little evidence for a causal role of NAFLD, T2D, obesity and glycemic traits in the regulation of *SMOC1* blood levels.

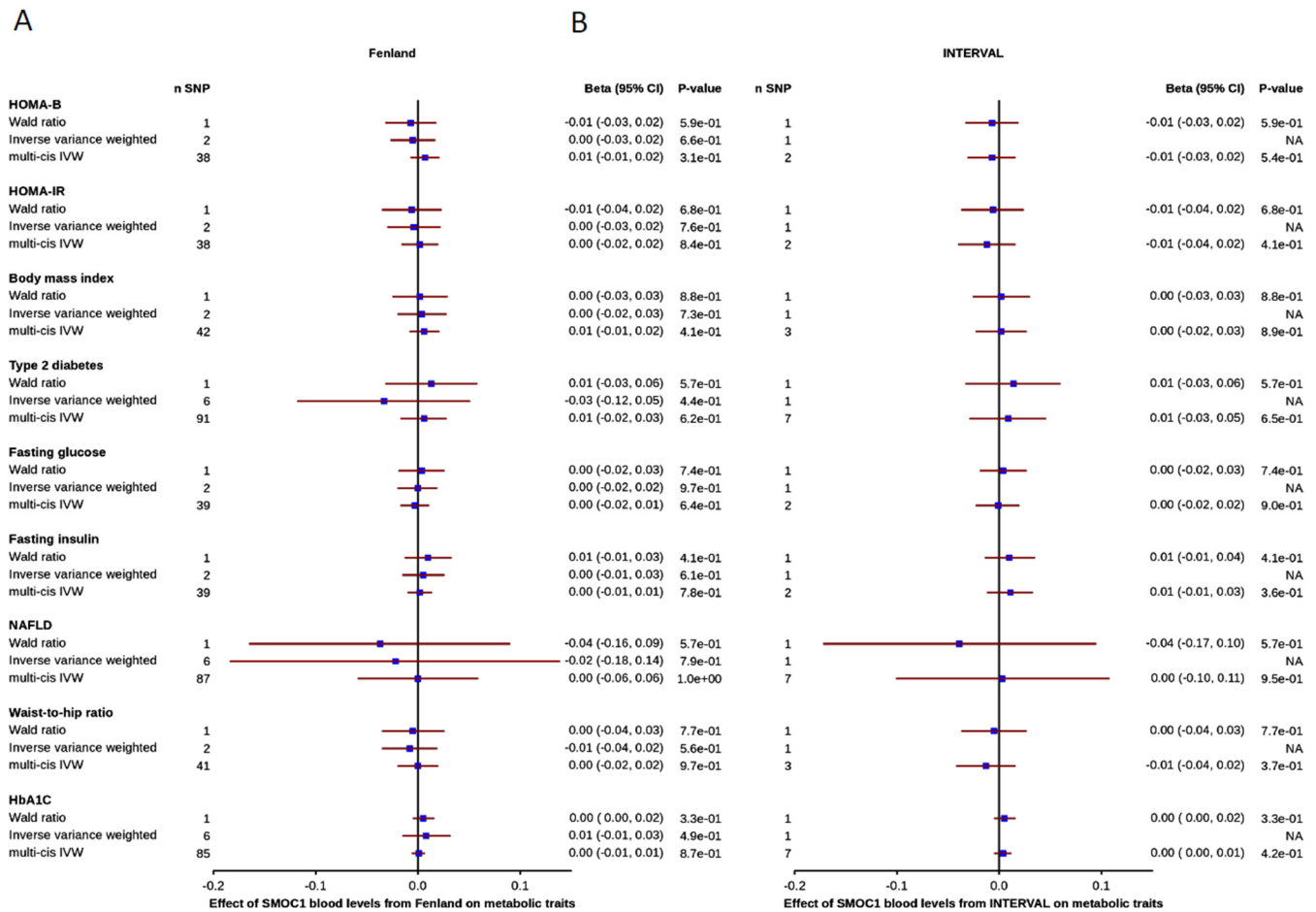


**Figure 2.** Impact of genetically predicted metabolic traits and diseases on blood *SMOC1* levels. Forest plot of Mendelian randomization estimated the effect of T2D, NAFLD, obesity and glycemic traits on blood *SMOC1* levels (panel A) and the effect of the waist-to-hip ratio on *SMOC1* levels across multiple MR methods (panel B).

### 3.3. Effect of Blood *SMOC1* Levels on Metabolic and Disease-Related Traits

The effect of blood *SMOC1* levels on glycemic traits are presented in Figure 3. First, the top cis-acting SNP from both INTERVAL and Fenland (rs1958078) was strong ( $R^2 = 0.05$ , F-statistic = 540 for Fenland study;  $R^2 = 0.02$ , F-statistic = 55 for INTERVAL study). Power was sufficient to detect small effects (power for Fenland study: mean = 0.95, SD = 0.14; power for INTERVAL study: mean = 0.80, SD = 0.23) to detect a 0.1 effect size (beta). No significant associations between *SMOC1* and disease-related and metabolic traits were found. Second, for the pan analysis, genetic instrument selection resulted in six SNPs for Fenland study and one SNP for INTERVAL study. Therefore, the INTERVAL study did not have enough instrument to carry this analysis. The use of additional genetic variants increased the exposure’s variance explained ( $R^2 = 0.078$  for Fenland study). Power using this method was higher (mean = 0.97, SD = 0.08) consistent with the use of more genetic instruments. Despite high power, IVW and robust MR methods for all exposure-outcome associations were non-significant. Third, for multi-cis analysis, genetic instrument selection resulted in 97 SNPs for the Fenland study and 7 SNPs for the INTERVAL study. Theoretically, several cis-acting variants can independently affect protein expression. We modelled this biological phenomenon by correcting for the LD matrix, which can result in a more precise estimate than the sole use of cis-acting sentinel SNP in QTL analysis [32]. The mean standard error across all evaluated associations for the multi-cis method was lower (0.012) compared to the sentinel SNP method (0.18). Despite increased precision,

no association was significant. Altogether, we found no evidence for a causal association between genetically predicted blood *SMOC1* levels and NAFLD, T2D, obesity as well as glycemic traits across all methods.



**Figure 3.** Impact of genetically predicted *SMOC1* levels on metabolic traits and diseases. Effect sizes and 95% confidence intervals for *SMOC1* SNPs derived from two GWAS studies (panel A) Fenland and (panel B) INTERVAL study. The blue cubes at each plot represent the Wald ratio estimate, inverse variance weighted (IVW) and multi cis IVW with 95% confidence interval.

### 3.4. Causal Effects of *SMOC1* Levels across the Human Phenome

We searched the top cis-acting SNP from both INTERVAL and Fenland studies (rs1958078) in the Phenoscanner [33,34] to discover new traits that may be associated with genetically predicted blood *SMOC1* levels. First, searching the eQTL database, rs1958078 was only and specifically associated at genome-wide significance with *SMOC1* gene level in the liver ( $p = 2 \times 10^{-22}$ ). Therefore, rs1958078 is a specific genetic instrument for *SMOC1* gene expression in the liver. This provides additional evidence that changing *SMOC1* levels in the liver might not result in improved glycemic control in humans. Second, searching the pQTL database, no other protein was associated with the rs1958078 SNP or its proxies ( $R^2 > 0.8$ ). This provides further evidence that the genetic instrument is specific to *SMOC1* and unlikely to be pleiotropic. Third, searching the GWAS database, we found that rs1958075 was associated with mean corpuscular hemoglobin (i.e., the average quantity of hemoglobin present in red blood cells) and red cell distribution width (i.e., the variation in red blood cell volume and size). Using cis-MR analysis, the level of *SMOC1* protein increased mean corpuscular hemoglobin and *SMOC1* level decreased red cell distribution width consistently in both Fenland and INTERVAL's studies. This indicates that elevating *SMOC1* plasma levels may be linked with blood-cell health outcomes. Altogether, our

search in the Phenoscanner confirmed the validity of our genetic instrument and that genetically predicted levels of the *SMOC1* hepatokine do not appear to influence metabolic traits but that could be linked to other blood cells-related traits.

#### 4. Discussion

To better understand the pathogenesis of metabolic diseases and to identify safe and effective prevention and treatment strategies, it is crucial to identify factors causally related to disease processes. Montgomery et al. recently proposed the hepatokine *SMOC1* as a potential therapeutic target for glycemic control based on their finding that *SMOC1* infusions and adenoviral-mediated hepatic *SMOC1* overexpression exerted a durable effect on glycemic control in mice [4]. *SMOC1*-increasing strategies also lead to improvements in hepatic insulin sensitivity in pre-clinical models. The authors concluded that *SMOC1* could represent a new therapeutic target that might be efficacious for treating T2D by improving glycemic control. In people with insulin resistance and obesity, *SMOC1* blood levels were statistically significantly lower than in people without these diseases. A recent study indicated that *SMOC1* is a thrombin-activating protein that makes a significant contribution to the pathophysiological changes in platelet function associated with type 2 diabetes [35]. *SMOC1* levels are correlated with the product of glucose and insulin measured during an oral glucose tolerance test [4]. These associations, however, do not provide evidence for a causal role of *SMOC1* in glucose-insulin homeostasis in humans.

The genomics revolution offers countless new opportunities to perform MR studies aimed at establishing causal relationships between a wide range of exposures and outcomes. Here, we performed bidirectional MR analyses to explore the causal associations of blood *SMOC1* levels with T2D, NAFLD, and glycemic traits. We found no evidence that NAFLD, T2D, and glycemic traits causally influenced *SMOC1* levels. We also did not find evidence that genetically predicted blood *SMOC1* levels were associated with T2D, NAFLD, and glycemic traits.

In comparison to observational analyses, the use of MR study design reduces the risk of bias from reverse causality and confounding factors. The large number of genetic variants with a strong effect on *SMOC1* levels and metabolic traits and diseases under investigation is a significant strength of our study. We also found no differences in *SMOC1* gene expression levels across stages of NAFLD. However, it should be kept in mind that the number of controls without NAFLD was rather limited ( $n = 10$ ).

In conclusion, these results contrast to those previously reported in mouse models and suggest that the hepatokine *SMOC1* might not have a causal role in the pathophysiology of human metabolic traits and diseases. *SMOC1* does not appear to represent an effective target to improve glycemic control in humans, although future randomized controlled trials in humans will be required to confirm this. However, these results suggest that blood factors regulating metabolism in rodents may not always translate to human biology.

**Author Contributions:** Conceptualization, B.J.A.; Methodology, N.G., J.B., É.G. and B.J.A.; Formal analysis, N.G., E.G. and J.B.; Writing—original draft preparation, N.G., E.G. and B.J.A.; Writing—review and editing, E.G., J.B., N.P., P.L.M., H.D.M. and A.G.; Supervision, B.J.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** E.G. holds a master's research award from the *Fonds de recherche du Québec: Santé*. (FRQS). B.J.A. holds a junior scholar award from the FRQS. This research received no external funding.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study, due to the fact that all the studies used for this manuscript were in the public domain.

**Informed Consent Statement:** All GWAS summary statistics used in this paper are publicly available and informed consent was obtained for all of the studies that were used.

**Acknowledgments:** We would like to thank all study participants as well as all investigators of the studies that were used throughout the course of this investigation.

**Conflicts of Interest:** B.J.A. is a consultant for Novartis and Silence Therapeutics and has received research contracts from Pfizer, Ionis Pharmaceuticals and Silence Therapeutics.

## References

1. Watt, M.J.; Miotto, P.M.; De Nardo, W.; Montgomery, M.K. The liver as an endocrine organ—Linking NAFLD and insulin resistance. *Endocr. Rev.* **2019**, *40*, 1367–1393. [[CrossRef](#)]
2. Meex, R.C.; Watt, M.J. Hepatokines: Linking nonalcoholic fatty liver disease and insulin resistance. *Nat. Rev. Endocrinol.* **2017**, *13*, 509–520. [[CrossRef](#)]
3. Meex, R.C.; Hoy, A.; Morris, A.; Brown, R.D.; Lo, J.C.; Burke, M.; Goode, R.; Kingwell, B.A.; Kraakman, M.; Febbraio, M.A.; et al. Fetuin B is a secreted hepatocyte factor linking steatosis to impaired glucose metabolism. *Cell Metab.* **2015**, *22*, 1078–1089. [[CrossRef](#)]
4. Montgomery, M.K.; Bayliss, J.; Devereux, C.; Bezawork-Geleta, A.; Roberts, D.; Huang, C.; Schittenhelm, R.B.; Ryan, A.; Townley, S.L.; Selth, L.A.; et al. SMOC1 is a glucose-responsive hepatokine and therapeutic target for glycemic control. *Sci. Transl. Med.* **2020**, *12*. [[CrossRef](#)]
5. Burgess, S.; Zuber, V. Risk Factors Mediating the Effect of Body-Mass Index and Waist-to-Hip 2 Ratio on Cardiovascular Outcomes: Mendelian Randomization Analysis. *Int. J. Obes.* **2021**, *45*, 1428–1438.
6. Jensen-Cody, S.O.; Potthoff, M.J. Hepatokines and metabolism: Deciphering communication from the liver. *Mol. Metab.* **2020**, *44*, 101138. [[CrossRef](#)] [[PubMed](#)]
7. Vannahme, C.; Smyth, N.; Miosge, N.; Gösling, S.; Frie, C.; Paulsson, M.; Maurer, P.; Hartmann, U. Characterization of SMOC-1, a novel modular calcium-binding protein in basement membranes. *J. Biol. Chem.* **2002**, *277*, 37977–37986. [[CrossRef](#)] [[PubMed](#)]
8. Bellenguez, C.; Strange, A.; Freeman, C.; Donnelly, P.; Spencer, C.C. A robust clustering algorithm for identifying problematic samples in genome-wide association studies. *Bioinformatics* **2012**, *28*, 134–135. [[CrossRef](#)] [[PubMed](#)]
9. Ruth, K.S.; Day, F.R.; Tyrrell, J.; Thompson, D.J.; Wood, A.R.; Mahajan, A.; Beaumont, B.N.; Wittemans, L.; Martin, S.; Busch, A.S. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat. Med.* **2020**, *26*, 252–258. [[CrossRef](#)]
10. Ding, E.; Song, Y.; Manson, J.E.; Hunter, D.J.; Lee, C.C.; Rifai, N.; Buring, J.E.; Gaziano, J.M.; Liu, S. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *N. Engl. J. Med.* **2009**, *361*, 1152–1163. [[CrossRef](#)]
11. Perry, J.R.; Weedon, M.N.; Langenberg, C.; Jackson, A.U.; Lyssenko, V.; Sparsø, T.; Thorleifsson, G.; Grallert, H.; Ferrucci, L.; Maggio, M.; et al. Genetic evidence that raised sex hormone binding globulin (SHBG) levels reduce the risk of type 2 diabetes. *Hum. Mol. Genet.* **2010**, *19*, 535–544. [[CrossRef](#)]
12. Govaere, O.; Cockell, S.; Tiniakos, D.; Queen, R.; Younes, R.; Vacca, M.; Alexander, L.; Ravaioli, F.; Palmer, J.; Petta, S.; et al. Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis. *Sci. Transl. Med.* **2020**, *12*, 572. [[CrossRef](#)]
13. Robinson, M.D.; Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **2010**, *11*, R25. [[CrossRef](#)]
14. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [[CrossRef](#)]
15. Sun, B.; Maranville, J.C.; Peters, J.; Stacey, D.; Staley, J.; Blackshaw, J.; Burgess, S.; Jiang, T.; Paige, E.; Surendran, P.; et al. Genomic atlas of the human plasma proteome. *Nature* **2018**, *558*, 73–79. [[CrossRef](#)]
16. Pietzner, M.; Wheeler, E.; Carrasco-Zanini, J.; Raffler, J.; Kerrison, N.D.; Oerton, E.; Auyeung, V.P.W.; Luan, J.; Finan, C.; Casas, J.P.; et al. Genetic architecture of host proteins involved in SARS-CoV-2 infection. *Nat. Commun.* **2020**, *11*, 6397. [[CrossRef](#)] [[PubMed](#)]
17. Ghodsian, N.; Abner, E.; Emdin, C.A.; Gobeil, É.; Taba, N.; Haas, M.E.; Perrot, N.; Manikpurage, H.D.; Gagnon, É.; Bourgault, J.; et al. Electronic health record-based genome-wide meta-analysis provides insights on the genetic architecture of non-alcoholic fatty liver disease. *Cell Rep. Med.* **2021**, *2*, 100437. [[CrossRef](#)]
18. Mahajan, A.; Taliun, D.; Thurner, M.; Robertson, N.R.; Torres, J.M.; Rayner, N.W.; Payne, A.J.; Steinthorsdottir, V.; Scott, R.A.; Grarup, N.; et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* **2018**, *50*, 1505–1513. [[CrossRef](#)] [[PubMed](#)]
19. Locke, A.E.; Kahali, B.; Berndt, S.I.; Justice, A.E.; Pers, T.H.; Day, F.R.; Powell, C.; Vedantam, S.; Buchkovich, M.L.; Yang, J.; et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* **2015**, *518*, 197–206. [[CrossRef](#)]
20. Shungin, D.; Winkler, T.W.; Croteau-Chonka, D.C.; Ferreira, T.; Locke, A.E.; Mägi, R.; Strawbridge, R.J.; Pers, T.H.; Fischer, K.; Justice, A.E.; et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **2015**, *518*, 187–196. [[CrossRef](#)] [[PubMed](#)]
21. Manning, A.K.; DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium; Hivert, M.-F.; Scott, A.R.; Grimsby, J.L.; Bouatia-Naji, N.; Chen, H.; Rybin, D.; Liu, C.-T.; Bielak, L.F.; et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat. Genet.* **2012**, *44*, 659–669. [[CrossRef](#)]
22. Dupuis, J.; Langenberg, C.; Prokopenko, I.; Saxena, R.; Soranzo, N.; Jackson, A.U.; Wheeler, E.; Glazer, N.L.; Bouatia-Naji, N.; Gloyn, A.L.; et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **2010**, *42*, 105–116. [[CrossRef](#)] [[PubMed](#)]



23. Soranzo, N.; Sanna, S.; Wheeler, E.; Gieger, C.; Radke, D.; Dupuis, J.; Bouatia-Naji, N.; Langenberg, C.; Prokopenko, I.; Stolerman, E.; et al. Common variants at 10 genomic loci influence hemoglobin A1C levels via glycemetic and nonglycemetic pathways. *Diabetes* **2010**, *59*, 3229–3239. [[CrossRef](#)]
24. Zhao, Q.; Chen, Y.; Wang, J.; Small, D.S. Powerful genome-wide design and robust statistical inference in two-sample summary-data Mendelian randomization. *arXiv* **2018**, arXiv:1804.07371.
25. Slob, E.A.; Burgess, S. A comparison of robust Mendelian randomization methods using summary data. *Genet. Epidemiol.* **2020**, *44*, 313–329. [[CrossRef](#)]
26. Bowden, J.; Smith, G.D.; Burgess, S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **2015**, *44*, 512–525. [[CrossRef](#)] [[PubMed](#)]
27. Verbanck, M.; Chen, C.-Y.; Neale, B.; Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **2018**, *50*, 693–698. [[CrossRef](#)]
28. Burgess, S.; Thompson, S.G.; Collaboration, C.C.G. Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiol.* **2011**, *40*, 755–764. [[CrossRef](#)] [[PubMed](#)]
29. Pierce, B.L.; Ahsan, H.; VanderWeele, T.J. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int. J. Epidemiol.* **2011**, *40*, 740–752. [[CrossRef](#)]
30. Burgess, S.; Bowden, J.; Fall, T.; Ingelsson, E.; Thompson, S.G. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiology* **2017**, *28*, 30. [[CrossRef](#)]
31. Burgess, S.; Zuber, V.; Valdes-Marquez, E.; Sun, B.B.; Hopewell, J.C. Mendelian randomization with fine-mapped genetic data: Choosing from large numbers of correlated instrumental variables. *Genet. Epidemiol.* **2017**, *41*, 714–725. [[CrossRef](#)] [[PubMed](#)]
32. Gkatzionis, A.; Burgess, S.; Newcombe, P.J. Statistical Methods for cis-Mendelian Randomization. *arXiv* **2021**, arXiv:2101.04081.
33. Staley, J.R.; Blackshaw, J.; Kamat, M.A.; Ellis, S.; Surendran, P.; Sun, B.B.; Paul, D.S.; Freitag, D.; Burgess, S.; Danesh, J.; et al. PhenoScanner: A database of human genotype–phenotype associations. *Bioinformatics* **2016**, *32*, 3207–3209. [[CrossRef](#)]
34. Kamat, M.A.; Blackshaw, J.A.; Young, R.; Surendran, P.; Burgess, S.; Danesh, J.; Butterworth, A.S.; Staley, J.R. PhenoScanner V2: An expanded tool for searching human genotype–phenotype associations. *Bioinformatics* **2019**, *35*, 4851–4853. [[CrossRef](#)] [[PubMed](#)]
35. Lagos, F.A.D.; Elgheznawy, A.; Kyselova, A.; Zu Heringdorf, D.M.; Ratiu, C.; Randriamboavonjy, V.; Mann, A.W.; Fisslthaler, B.; Siragusa, M.; Fleming, I. Secreted modular calcium-binding protein 1 binds and activates thrombin to account for platelet hyperreactivity in diabetes. *Blood* **2021**, *137*, 1641–1651. [[CrossRef](#)] [[PubMed](#)]