European Heart Journal Supplements (2020) **22** (Supplement C), C34–C45
The Heart of the Matter
doi:10.1093/eurheartj/suaa008

ESC
European Society
of Cardiology

Bayesian network analysis of plasma microRNA sequencing data in patients with venous thrombosis

Florian Thibord[1,2], Gaëlle Munsch[1], Claire Perret[3], Pierre Suchon[4], Maguelonne Roux[3], Manal Ibrahim-Kosta[4,5], Louisa Goumidi[5], Jean-François Deleuze[6,7], Pierre-Emmanuel Morange[4,5†], and David-Alexandre Trégouët[1*†]; on behalf of the GENMED Consortium

[1]*Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR_S) 1219, Bordeaux Population Health Research Center, University of Bordeaux, 146 rue Léo Saignat, Bordeaux 33076, France;*
[2]*Pierre Louis Doctoral School of Public Health, Sorbonne-Université, 15 rue de l'école de médecine, Paris 75006, France;*
[3]*Sorbonne Universités, Université Pierre et Marie Curie (UPMC Univ Paris 06), INSERM UMR_S 1166, 91 Boulevard de l'Hôpital, Paris 75013, France;*
[4]*Laboratory of Haematology, La Timone Hospital, 278 rue Saint Pierre, Marseille 13385, France;*
[5]*INSERM UMR_S 1062, Nutrition Obesity and Risk of Thrombosis, Center for CardioVascular and Nutrition research (C2VN), Aix-Marseille University, 278 rue Saint Pierre, Marseille 13385, France;*
[6]*Centre National de Recherche en Génomique Humaine, Direction de la Recherche Fondamentale, CEA, 2 rue Gaston Crémieux, Evry 91057, France; and*
[7]*CEPH, Fondation Jean Dausset, 27 rue Juliette Dodu, Paris 75010, France*

KEYWORDS
Venous thrombosis;
plasma miRNA;
Next generation sequencing;
Biomarkers;
Genome Wide
Association Study

MicroRNAs (miRNAs) are small regulatory RNAs participating to several biological processes and known to be involved in various pathologies. Measurable in body fluids, miRNAs have been proposed to serve as efficient biomarkers for diseases and/or associated traits. Here, we performed a next-generation-sequencing based profiling of plasma miRNAs in 344 patients with venous thrombosis (VT) and assessed the association of plasma miRNA levels with several haemostatic traits and the risk of VT recurrence. Among the most significant findings, we detected an association between hsa-miR-199b-3p and haematocrit levels ($P = 0.0016$), these two markers having both been independently reported to associate with VT risk. We also observed suggestive evidence for association of hsa-miR-370-3p ($P = 0.019$), hsa-miR-27b-3p ($P = 0.016$) and hsa-miR-222-3p ($P = 0.049$) with VT recurrence, the observations at the latter two miRNAs confirming the recent findings of Wang *et al*. Besides, by conducting Genome-Wide Association Studies on miRNA levels and meta-analyzing our results with some publicly available, we identified 21 new associations of single nucleotide polymorphisms with plasma miRNA levels at the statistical significance threshold of $P < 5 \times 10^{-8}$, some of these associations pertaining to thrombosis associated mechanisms. In conclusion, this study provides novel data about the impact of miRNAs' variability in haemostasis and new arguments supporting the association of few miRNAs with the risk of recurrence in patients with venous thrombosis.

---

[*] Corresponding author. Tel: +33 5 4730 4254, Email: david-alexandre.tregouet@u-bordeaux.fr; david-alexandre.tregouet@inserm.fr
[†] These authors contributed equally to this study.

PALABRAS CLAVE
Trombosis venosa;
Plasma miRNA;
Secuenciación de última generación;
Biomarcadores;
Genoma completo;
Estudio de asociación

Los micro-ARN (miARN) son pequeñas moléculas de ARN reguladoras que participan en varios procesos biológicos y están implicados en diversas patologías. Mensurables en los líquidos corporales, se ha planteado que los miARN pueden ser biomarcadores eficaces para el diagnóstico de enfermedades y/o características asociadas. Aquí hemos llevado a cabo un análisis de miARN plasmático con tecnología de secuenciación de última generación en 344 pacientes con trombosis venosa (TV) y hemos evaluado la asociación de los niveles de miARN con distintas características hemostáticas y el riesgo de recidiva de TV. Entre los hallazgos más significativos, hemos detectado una asociación entre hsa-miR-199b-3p y los niveles de hematocritos ($p = 0{,}0016$); dos marcadores que se habían asociado de forma independiente con el riesgo de sufrir TV. Asimismo, hemos observado una evidencia indicativa de asociación entre hsa-miR-370-3p ($p = 0{,}019$), hsa-miR-27b-3p ($p = 0{,}016$) y hsa-miR-222-3p ($p = 0{,}049$) y la recidiva de TV; los resultados los dos últimos miARN confirman los hallazgos recientes de Wang *et al.* (Clin Epigenetics, 2019). Además, al efectuar estudios de asociación del genoma completo sobre los niveles de miARN y al metaanalizar nuestros resultados con otros disponibles públicamente, hemos identificado 21 asociaciones nuevas de polimorfismos de un solo nucleótido (PSN) con niveles de miARN plasmático con un umbral de significación estadística de $p < 5 \times 10^{-8}$; algunas de estas asociaciones pertenecen a los mecanismos patogénicos de la trombosis.

Como conclusión, en este estudio se proporcionan nuevos datos sobre el impacto de la variabilidad de miARN en la hemostasia y nuevos argumentos que apoyan la asociación de algunas secuencias de miARN con el riesgo de recidiva en pacientes con trombosis venosa.

关键词
静脉血栓形成;
血浆miRNA;
下一代测序;
生物标记;
基因组广泛;
关联研究

微小 RNA (miRNA) 是参与多种生物学过程的小型调节性 RNA，已知参与各种病理过程。在体液中可测量的 miRNA 已被提议为疾病和/或相关性状的有效生物标记。我们在此对 344 名静脉血栓形成 (VT) 患者的血浆 miRNA 进行了基于下一代测序的分析，并评估了血浆 miRNA 水平与几种止血性状和 VT 复发风险之间的关系。我们的重大发现之一是，我们检测到 hsa-miR-199b-3p 与血细胞比容水平之间存在关联 ($p = 0.0016$)，这两个标志物均已被独立报道与 VT风险相关。我们还观察到了提示性的证据， 表明 hsa-miR-370-3p ($p = 0.019$)、hsa-miR-27b-3p ($p = 0.016$) 和 hsa-miR-222-3p ($p = 0.049$) 与 VT 复发相关， 后两个 miRNA 的观察结果证实了 Wang □□的最新发现。（临床表观遗传学 (Clin Epigenetics)，2019 年）。此外，通过对 miRNA 水平进行基因组广泛关联研究并通过一些可获得的公开结果对我们的结果进行荟萃分析， 我们在 $p < 5 \times 10^{8}$ 的统计学显著性阈值下发现了 21 个 SNP 与血浆 miRNA 水平的新关联，其中一些关联与血栓形成的相关机制有关。

总而言之， 这项研究提供了有关 miRNA 变异性在止血方面的影响的新数据， 并为少数 miRNA 与静脉血栓形成患者复发风险的相关性提供了新论据。

Introduction

Venous thrombosis (VT), including deep vein thrombosis (DVT) and pulmonary embolism (PE), affects about 1 200 000 individuals each year in Europe and is thus the third most common cardiovascular disease after coronary artery disease and stroke.[1] It is a severe disorder that leaves many patients (25-50%) with a debilitating post-thrombotic syndrome[2] and whose PE manifestation kills many of them (6% acute and 20% after 1 year).[3] About 50% of VT are unprovoked, i.e. they occur without clear external factors like surgery, trauma, immobilization, hormone use or cancer. The annual recurrent rate is ∼6% and about 25% of patients with unprovoked VT will face a recurrent event after a 6-month course of anticoagulant treatment.[4] Thus, the secondary prevention of VT in this specific population group of patients with a first unprovoked VT is a major health issue.

There is an urgent need to better understand the pathophysiological mechanisms leading to VT in order to develop targeted therapeutic and preventative strategies to save lives, improve quality of life and reduce healthcare costs. Effective preventative options are available in the form of anticoagulant treatments, but these are associated with major bleeding complications. There are unmet needs to develop predictive biomarkers with high sensitivity and specificity for accurate identification of patients who will develop a recurrence, to avoid unacceptably high risk of bleeding complications in patients at low risk of recurrence. Indeed, preventing thrombosis without inducing bleeding is the holy grail of anticoagulant therapy. Currently, there are no commercially available anticoagulants that achieve this.

Predicting the risk of recurrence and discriminating between fatal (PE) and non-fatal (DVT) events in unprovoked VT patients remain challenging. There is so far no established biomarkers that serve these aims, even if D-dimers measurement has been proposed[5] but lacks specificity. We here propose a comprehensive microRNA (miRNA) profiling from plasma samples of VT patients aimed at discovering miRNA-derived biomarkers discriminating between PE and DVT and associated with VT recurrence. MicroRNAs represent a class of small (~22 nucleotides) non-coding RNAs that participate in genes post-transcriptional regulation.[6] It is now well-established that miRNAs are involved in the development of human diseases, in particular, cardiovascular ones.[7] Several genes participating to thrombosis associated mechanisms have already been suspected to be subject to miRNA regulation.[8-11] So far, epidemiological studies looking for association of plasma miRNAs with VT outcomes are still sparse. Using plasma samples of 20 VT cases and 20 healthy individuals, Starikova *et al.*[12] assessed the association of 97 miRNAs with VT risk among which 9 were found significantly ($P < 0.05$) associated with the outcome. As for Wang *et al.*,[13] by looking for the association of 110 miRNAs with the risk of VT recurrence in plasma samples of 39 cases and 39 controls, 12 miRNAs were identified. None of these observations, which were obtained on miRNA data profiled using RT-qPCR techniques, have yet been replicated.

Briefly, we here performed plasma miRNA profiling in 391 VT patients using a next-generation sequencing technology and assessed the association of identified miRNAs with several haemostatic traits and VT associated clinical outcomes. Association analyses were conducted using an original Bayesian network (BN) inference strategy aimed at identifying miRNAs with the highest abilities to serve as relevant biomarkers. In addition, we integrated genome-wide genotype data with miRNA expression levels in order to identify miRNAs that are under strong genetic control.

Methods

The MARTHA microRNA sequencing study

The MARseille THrombosis Association project refers to a collection of VT patients recruited at the La Timone Hospital in Marseille, France, initially between 1994 and 2005 and further extended over the 2010–12 period. Detailed description of this collection has already been previously provided.[14]

The present study relies on a subsample of 391 VT patients that had been previously genotyped for genome-wide polymorphisms using dedicated genotyping array[15,16] and with available plasma samples. For each sample, total RNA was extracted from 400 μL citrate plasma sample using miRNeasy Serum/Plasma kit from Qiagen. From 6 μL of total RNA, plasma miRNA libraries were then prepared with NEBNext Multiplex Small RNA Library Prep Set for Illumina. The manufacturer's protocol was followed, with an optimized size selection method via Ampure XP beads, a specific dilution of adapters to 1/10, and 15 cycles of PCR amplification, using adapter sequences GATC GG AAGAGCACACGTCTGAACTCCAGTCAC and CGACAGGTTCAG AGTTCTACAGTCCGACGATC for 3' and 5' ends, respectively. Detailed characteristics of the experimental protocol for libraries preparation and sequencing have already been described.[17]

MicroRNA alignment and quantification processes

Sequenced data were processed with the bioinformatic OptimiR pipeline[17] in order to detect and quantify miRNAs. Briefly, OptimiR aligned miRNAs to a library composed of mature miRNA references sequences from miRBase 21.[18] For miRNA integrating genetic variants in their sequence (called polymiRs), the reference library was upgraded by OptimiR with sequences integrating alternative alleles. Ambiguous alignments were resolved using a scoring algorithm that keeps only the most likely alignment while considering the frequent post-transcriptional modifications that miRNAs can undergo.[19] Reads aligned on polymiRs were kept if they were consistent with the sample's genotype, otherwise, they were discarded.[17]

From the resulting miRNA abundances, we performed several quality assessments in order to discard unreliable data. First, samples that were poorly sequenced, i.e. with <100 000 reads aligned, were discarded ($n = 3$) as well as samples identified to be haemolyzed ($n = 34$). The degree of haemolysis was determined based on the optical density at 414 nm, and values exceeding 0.2 were defined as haemolyzed samples.[20] Finally, in order to retain only highly expressed miRNAs, we kept only those with at least five counts in at least 75% of the remaining samples.

Abundances were then normalized using the rlog method from the DESeq2 R library.[21] This normalization process takes into account differences in library sizes due to library preparation and sequencing protocols, and stabilize variance across miRNAs and samples to respect homoscedasticity constraints for further analysis. Principal component analysis (PCA) was applied to normalized abundances in order to identify individuals with outliers miRNA profiles. Individuals deviating by 3 SD from the centres of the first four PCAs ($n = 10$) were further excluded from downstream analyses, leaving 344 individuals for BN and association analyses.

Bayesian network analysis

A BN is a probabilistic directed acyclic graphical model that represents relationships among a large number of variables (here mainly miRNAs) with the aim of modelling the dependencies/interactions and conditional independencies between variables.[22,23] Generally, any BN is defined by a directed acyclic graph structure $G = (V,E)$ where V is the set of variables and E the set of edges representing the directional relationships between variables and P a joint probability distribution of the variables in the network. Three types of nodes can be identified in a given BN: the root nodes that are variables found to influence several other variables but are not themselves influenced by any other variables, the internal nodes that are both influenced by and modulate other variables, and finally terminal nodes that are variables that are not identified as influencing others (see *Figure 1*). Any variable influencing another variable in the network is referred to as a parental node for this later variable. In the following, we will mainly

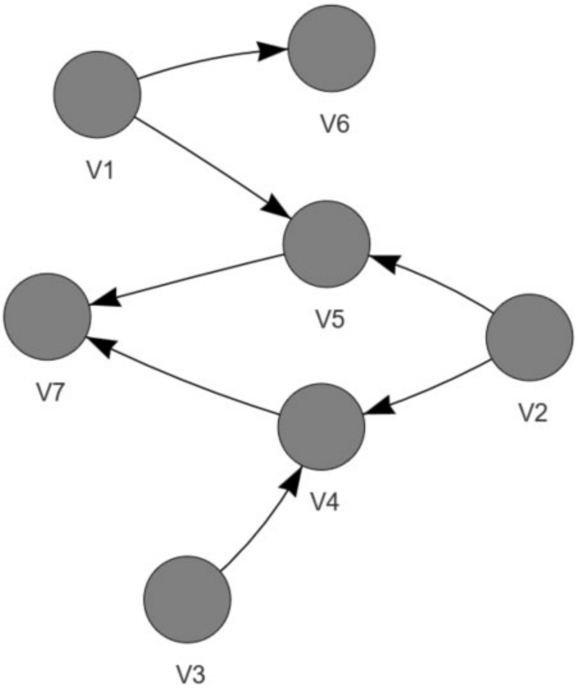

Figure 1 A Bayesian network example. In this illustrative BN example, variables V1, V2, and V3 are root nodes, V4 and V5 are internal nodes, and V6 and V7 are terminal nodes. V3 is also a parental node for V4 which is itself a parental node for V7.

focus on terminal nodes assuming that such nodes, as integrating the cumulative upstream effects of other variables, would serve as more relevant and powerful endophenotypes to be tested in relation to some outcomes of interest. In that context, BN analysis can also be viewed as a data reduction technique since, instead of testing the association of all initial variables with a given outcome, only the terminal nodes will be tested for association, reducing then the multiple testing burden. In this article, BNs will be constructed with the «bnlearn» package[24] that implements the relatively fast *tabu search* algorithm handling both discrete and continuous variables. In the current application, BNs will be created from all expressed miRNAs but also with the age and sex variables. These two latter variables have been shown to have strong influence on circulating miRNA levels[25,26] and their integration in the BN analysis can then add information to more efficiently model the dependencies and conditional independence between some miRNAs.

Because *tabu search* is a greedy search algorithm, it may end up into a local optimum. To overcome such situation and to assess the stability of the BN analysis in identifying robust terminal nodes, we generated 2000 bootstrapped datasets composed of 95% of the initial samples and for each boostrapped datasets, we randomly shuffled the way the input variables were ordered in the initial dataset. For each shuffled bootstrapped dataset, a BN was constructed and the terminal nodes identified. After 2000 bootstrap, we calculated the number of times a given variable was identified as terminal node.

In order to assess whether the observed distribution of the number of terminal node's occurrences deviates from the null hypothesis of no correlation structure between miRNAs, a permutation strategy was adopted. For each permutation, we randomly selected at least 40 variables whose values were permuted between individuals in order to break down the original data correlation structure. We generated 2000 of such permuted datasets and constructed a BN on each of them. From these permuted BNs, we counted the maximum number of times a given variable (that could be any miRNA, age, or sex) was identified as a terminal node and used this maximum value as a cut-off to identify robust terminal miRNAs in the unpermuted analysis above.

Association analysis with haemostatic traits and clinical outcomes

Identified terminal miRNAs were tested for association with several haemostatic traits available in MARTHA participants (see *Table 1*). Association analyses were performed using linear regression model and adjusted for age, sex, anticoagulant therapy, and combined plasma levels of hsa-let-7d-5p, hsa-let-7g-5p and let-7i-5p measured by qPCR, which serve as a control reference of miRNA levels.[27] Individuals under anticoagulant therapy at the time of blood sampling were excluded for the analysis on protein C, protein S, and prothrombin time. For association testing, log-transformation was applied to the following variables: Activated Thrombin Generation Potential biomarkers (Endogenous Thrombin Potential, Lagtime), Partial Thromboplastin Time, Factor VIII, Homocystein, Plasminogen Activator Inhibitor-1, Tissue Factor Principal Inhibitor, and von Willebrand Factor.

Terminal miRNAs were also tested for association with the DVT vs. PE outcome using a logistic regression model, while a Cox model was used to assess their association with VT recurrence whose information was available in 228 patients only. For the latter analysis, we applied the Cox survival model with left truncature[28] and adjusted for age, sex, body mass index (BMI), and smoking. To address the multiple testing issue associated with the number of terminal miRNAs that will be tested for association with the phenotypes, we applied a Bonferroni correction based on the effective number of independent variables.[29]

Genome-wide miR-eQTL analysis

As MARTHA participants have been typed for high-density genotyping arrays and imputed for common polymorphisms available in the 1000 G reference panel, we performed genome-wide association study (GWAS) on each expressed miRNA for identifying miRNA expression quantitative trait loci (miR-eQTL) using the mach2QTL programme.[30] Analyses were performed under the assumption of additive genetic effects and adjusting for the following covariates: sex, age of blood collection, anticoagulant prescription, RT-qPCR measured hsa-let-7 combination,[27] and the four first principal genetic components retrieved from PCA analysis as previously described.[15,16] GWAS results were filtered out for variants with minor allele frequency lower than 0.05 and with imputation criterion r^2 below 0.4. Finally, we combined the results of our miR-eQTL analysis with those previously described by Nikpay *et al.*[31] and

Table 1 Characteristics of the MARTHA miRNA cohort

Variables	*N*	Mean ± SD[a]
Gender (male/female)	344	144/200
Age (years)	344	52.1 ± 14.5
Smoking (yes/no)	343	94/249
BMI (kg/m^2)	331	25.86 ± 4.62
Deep vein thrombosis/pulmonary embolism	344	259/85
Anticoagulant therapy (yes/no)	344	122/222
Antithrombin (IU/mL)	313	102.41 ± 11.59
Activated partial thromboplastin time (s)	341	33.42 ± 6.02
D-dimers (μg/mL)	184	0.39 ± 0.33[b]
FV (IU/mL)	150	109.21 ± 22.26
FVIII (IU/dL)	294	135.07 ± 48.31
FXI (IU/mL)	336	130.78 ± 31.99
Fibrinogen (g/L)	342	3.42 ± 0.66
Haematocrit (L/L)	343	0.42 ± 0.03
Homocysteine (μmol/L)	304	12.26 ± 5.65
Platelet count (G/L)	344	254.62 ± 64.91
Mean platelet volume (fL)	344	7.90 ± 0.77
Haemoglobin (g/dL)	344	140.42 ± 13.19
PAI-1 (UI/mL)	272	12.25 ± 13.44
Protein C (IU/mL)	318	99.55 ± 40.56
Protein S (IU/mL)	322	81.3 ± 27.49
TAFI (μg/mL)	336	15.27 ± 4.72
TFPI (ng/mL)	336	14.17 ± 6.84
vWF (IU/dL)	308	154.34 ± 67.74
Prothrombin time (%)	344	87.63 ± 27.95
Thrombin generation	193	
Endogeneous thrombin potential (nM·min)		1761.44 ± 280.31
Peak (nM)		340.35 ± 57.51
Lagtime (min)		3.34 ± 1.17
VT recurrence during follow-up (yes/no)	228	41/187

[a] Count data are shown for categorical variables, other reported values were mean ± standard deviation.

[b] In about 50% participants, D-dimers values were below the detection limit (0.22) and thus discarded. Mean and SD were then computed over all D-dimer values >0.22.

available at https://zenodo.org/record/2560974 in order to identify additional single nucleotide polymorphism (SNP) × miRNA associations. For this, a random-effect model-based meta-analysis was adopted as implemented in the GWAMA software.[32] SNP × miRNA associations were considered as *cis* effects when the SNP maps ± 1 Mb from the mature miRNA position. Otherwise, they were considered as *trans*. Any association with P-value $<3.2 \times 10^{-10}$ corresponding to the Bonferroni threshold corrected for the number of tested SNP × miRNA associations was considered as genome-wide significant. We also used a miRNA-wide threshold of $P<5 \times 10^{-8}$, the standard statistical threshold generally advocated in the context of a single GWAS, to identify additional suggestive associations.

Results

The MARTHA microRNA cohort

Detailed description of the clinical and biological characteristics of the 344 participants is shown in *Table 1*. Of note, 228 patients have been followed for the risk of recurrence for a mean time period of 11.4 ± 4.3 years. During this period, 41 patients experienced a new VT event.

After the application of the OptimiR workflow, 162 miRNAs were found expressed in the 344 MARTHA participants. Full miRNA data are provided in Supplementary material online, *Table S1*. The most expressed miRNA was the hsa-miR-122-5p (Supplementary material online, *Figure S1*), a miRNA known to be mainly expressed in liver and that was previously shown to be amongst the most abundant plasma miRNAs.[33] Additional highly expressed miRNAs were hsa-miR-486-5p, hsa-miR-92a-3p, and hsa-miR-451a (Supplementary material online, *Figure S1*). Of note, the 25 most expressed miRNAs accounted for >90% of all sequenced reads that were aligned to miRNA mature sequences.

BN analysis of microRNA data

Under the null hypothesis of no specific structure in the miRNA data, all miRNAs were identified as a terminal node at least once and, on average, a miRNA was found as a

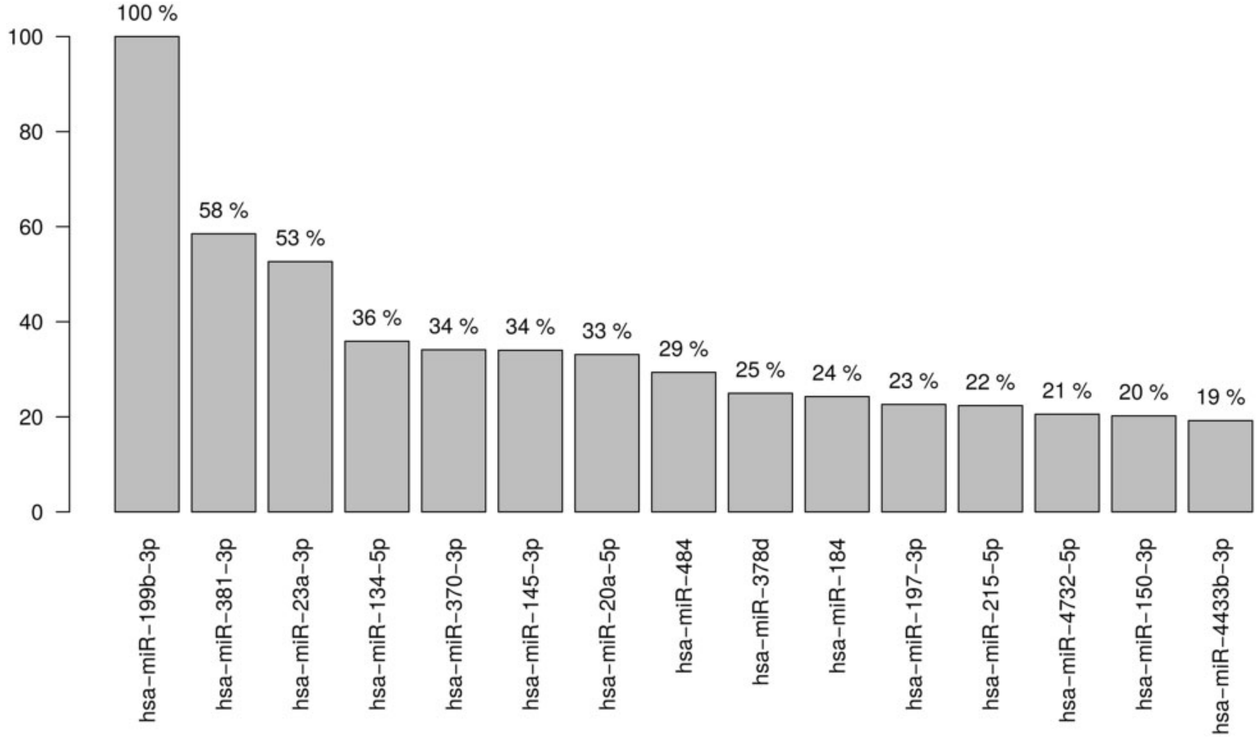

Figure 2 Percentage of significant terminal miRNAs found in 2000 bootstrapped Bayesian networks. The bootstrap BN analysis identified 15 terminal miRNAs with an occurrence percentage over the significance threshold (18.3%) determined by the permutation analysis.

terminal node in 6.3% ± 3.5 of the permuted BNs, with a maximum of 18.3%. Using the latter threshold, the bootstrap BN analysis identified 15 terminal miRNAs and the number of times each of them was found as a terminal node in boostrapped BNs is shown in *Figure 2*.

Association of microRNAs' levels with VT-associated biological and clinical traits

The application of the Li and Ji multiple testing procedure[29] estimated the number of effective independent terminal miRNAs as 14, leading to an adapted Bonferroni threshold of 3.6×10^{-3}. At this statistical level, only one association between terminal miRNAs and haemostatic traits was detected. Plasma levels of hsa-miR-199b-3p were negatively correlated ($\rho = -0.17$, $P = 0.0016$) with haematocrit levels. Interestingly, this miRNA has recently been reported to associate with VT risk[12] whose association with haematocrit levels have already been described.[34,35] The full results of the scan for association between miRNAs and haemostatic traits are given in Supplementary material online, *Table S2*.

Of note, the strongest association of terminal miRNAs with recurrence risk was observed for hsa-miR-370-3p [HR = 1.77 (1.09-2.88), $P = 0.019$], this miRNA being also the terminal miRNA that discriminated the most between DVT and PE [OR for PE = 0.72 (0.49-1.05), $P = 0.090$] (*Table 2*). Of interest, one of our terminal miRNAs, hsa-miR-197-3, was reported to associate with VT recurrence in Wang *et al.*[13] However, we did not observe here such trend for association [HR = 0.78 (0.35-1.76), $P = 0.55$]. Nevertheless, among the nine additional miRNAs reported in Wang *et al.* and also expressed in MARTHA, we found two with a suggestive association with VT recurrence: hsa-miR-27b-3p [HR = 0.4 (0.2-0.79), $P = 0.016$] and hsa-miR-222-3p [HR = 1.76 (1.01-3.08), $P = 0.049$] (Supplementary material online, *Table S3*).

miR-eQTL analyses

At the pre-specified genome-wide statistical level of 3.2×10^{-10}, three SNP × miRNA associations, all *cis*, were identified in the MARTHA study (*Table 3*). These were observed for rs12473206 with hsa-miR-4433b-3p ($P = 8.12 \times 10^{-35}$), rs2127870 with hsa-miR-625-3p ($P = 9.57 \times 10^{-26}$), and rs140930133 with hsa-miR-941 ($P = 5.07 \times 10^{-15}$). The latter two have already been observed in whole blood[36] and adipose tissue.[37] Using a more liberal miRNA-wide threshold of $P = 5 \times 10^{-8}$, 10 additional suggestive associations, 1 in *cis* and 9 in *trans*, were observed (*Table 3*). Regional association plots and boxplot summarizing the genotype × miRNA associations at these 13 main candidates are shown in Supplementary material online.

Of note, the most significant association was observed between hsa-miR-4433b-3p and rs12473206, a variant located within the mature miRNA sequence. It can be speculated that this variant impacts the maturation process of the miRNA or its target spectrum, and thus influences its plasma expression levels. In addition, two SNPs with *cis* effects on miRNA levels (thereafter referred to as *cis* miSNPs) have been previously found to associate with levels of the protein encoded by the miRNA host gene. In whole blood, the miSNP rs2127870 was reported to

Table 2 Association of terminal miRNAs with VT outcomes in the MARTHA miRNA study

miRNA	VT recurrence		Pulmonary embolism vs. deep vein thrombosis	
	HR (95% CI)	P^a	OR (95% CI)	P^b
hsa-miR-370-3p	1.77 (1.09-2.88)	0.019	0.72 (0.49-1.05)	0.090
hsa-miR-184	0.53 (0.30-0.95)	0.024	1.23 (0.92-1.66)	0.153
hsa-miR-4732-5p	0.41 (0.18-0.92)	0.024	0.70 (0.39-1.22)	0.218
hsa-miR-4433b-3p	1.54 (1.04-2.29)	0.033	1.01 (0.75-1.36)	0.930
hsa-miR-215-5p	0.63 (0.37-1.09)	0.091	1.11 (0.73-1.67)	0.633
hsa-miR-134-5p	1.58 (0.85-2.91)	0.142	0.89 (0.57-1.39)	0.601
hsa-miR-381-3p	1.45 (0.83-2.56)	0.194	0.81 (0.53-1.23)	0.327
hsa-miR-145-3p	0.51 (0.15-1.76)	0.278	0.62 (0.24-1.56)	0.311
hsa-miR-23a-3p	0.67 (0.26-1.70)	0.393	1.00 (0.51-1.93)	0.999
hsa-miR-197-3p	0.78 (0.35-1.76)	0.555	1.41 (0.79-2.56)	0.251
hsa-miR-150-3p	1.23 (0.53-2.83)	0.629	0.90 (0.49-1.66)	0.743
hsa-miR-484	1.20 (0.56-2.59)	0.637	1.27 (0.69-2.38)	0.447
hsa-miR-199a-3p	0.80 (0.22-2.86)	0.726	1.17 (0.46-2.97)	0.746
hsa-miR-378d	0.81 (0.15-4.56)	0.812	0.41 (0.10-1.46)	0.184
hsa-miR-20a-5p	1.09 (0.40-2.95)	0.863	0.74 (0.36-1.52)	0.411

[a] *P*-values were obtained from the Likelihood Ratio test statistic associated with a Cox survival model adjusted for age, sex, BMI, and smoking.
[b] *P*-values obtained from a logistic model adjusted for age, sex, BMI, and smoking.

Table 3 Significant associations at the $5\cdot10^{-8}$ statistical level between SNPs and plasma miRNA levels in the MARTHA miRNA study

miRNA	miRNA host gene	Top SNP Associated	MAF	r^2	Chr	Distance to 5' miRNA	Effect (SD)	*P*-value	SNP Genomic Context
Cis associations									
hsa-miR-4433b-3p	Intergenic	rs12473206	0.23	0.99	2	−13	0.979 (0.080)	$8.12\cdot10^{-35}$	exonic_ncRNA (hsa-miR-4433b)
hsa-miR-625-3p	FUT8	rs2127870	0.27	0.99	14	141 025	0.533 (0.051)	$9.57\cdot10^{-26}$	Intergenic
hsa-miR-941	DNAJC5	rs140930133	0.19	0.97	20	8822	−0.349 (0.045)	$5.07\cdot10^{-15}$	Intronic (DNAJC5)
hsa-miR-432-5p	RTL1	rs201969986	0.29	0.95	14	177 423	−0.346 (0.063)	$3.31\cdot10^{-8}$	Intergenic
Trans associations									
hsa-miR-184		rs144867605	0.07	0.82	11	75 957 983	0.804 (0.134)	$2.02\cdot10^{-9}$	Intergenic
hsa-miR-654-5p		rs11109171	0.44	0.99	12	98 098 091	−0.246 (0.042)	$3.28\cdot10^{-9}$	Intergenic
hsa-miR-320c		rs10151482	0.06	0.93	14	41 934 917	0.427 (0.074)	$6.47\cdot10^{-9}$	Intergenic
hsa-miR-184		rs143007764	0.06	0.65	3	142 899 139	0.916 (0.161)	$1.14\cdot10^{-8}$	Intergenic
hsa-miR-1-3p		rs73245753	0.12	0.79	4	26 292 392	0.589 (0.105)	$2.31\ 10^{-8}$	Intergenic
hsa-miR-330-3p		rs1554362	0.45	0.82	2	101 221 457	−0.227 (0.041)	$2.81\cdot10^{-8}$	Intronic (LINC01849)
hsa-miR-582-3p		rs4522365	0.13	0.83	15	29 964 742	0.314 (0.057)	$2.91\cdot10^{-8}$	Intergenic
hsa-miR-4446-3p		chr12:95274192:I	0.09	0.61	12	95 274 192	−0.492 (0.089)	$3.07\cdot10^{-8}$	Intergenic
hsa-miR-320d		rs12800249	0.05	0.63	11	21 240 436	0.481 (0.088)	$4.33\cdot10^{-8}$	Intronic (NELL1)

MAF, minor allele frequency; r^2, imputation quality criterion.

influence FUT8 levels,[38] *FUT8* being the host gene for hsa-miR-625-3p. Similarly, the *DNAJC5* rs2427555 that is in very strong linkage disequilibrium (LD) with the miSNP rs140930133, we here found associated with plasma hsa-miR-941 levels, has been reported to influence the expression of *DNAJC5* in lymphoblastoid cells.[39] These observations are supportive elements for the observed miSNP associations and would suggest a joint regulation of hsa-miR-625-3p and hsa-miR-941 expressions with those of their host genes as already documented for several miRNAs.[40]

One *trans*-eQTL located in the long non-coding RNA (lncRNA) LINC01849 was associated with hsa-miR-330-3p. The identified *trans* miSNP, rs1554362, is also an eQTL for

Table 4 Association of SNPs with plasma miRNA levels identified in Nikpay *et al.*[31] that nominally replicated ($P < 0.05$) in MARTHA miRNA study

					NIKPAY ($N = 710$)				MARTHA ($n = 344$)				
miRNA	SNP	Chr	Position(bp)	EA	EAF	β	SE	P	EAF	R^2	β	SE	P^a
Cis associations													
miR-941	rs2427550	20	62547575	A	0.23	−0.157	0.023	3.96×10^{-11}	0.19	0.99	−0.339	0.044	5.76×10^{-15}
miR-584-5p	rs17795259	5	148416952	C	0.15	0.268	0.018	1.35×10^{-45}	0.15	0.99	0.213	0.043	4.82×10^{-7}
miR-4433b-5p	rs2059631	2	64574682	A	0.43	0.289	0.017	1.57×10^{-56}	0.45	1.00	0.129	0.029	4.96×10^{-6}
miR-139-3p	rs4944563	11	72316881	C	0.17	0.169	0.026	1.18×10^{-10}	0.14	1.00	0.182	0.042	6.82×10^{-6}
miR-181a-5p	rs74746864	1	199023240	G	0.11	0.175	0.025	4.12×10^{-12}	0.13	0.95	0.221	0.066	4.27×10^{-4}
miR-425-5p	rs7623513	3	142100428	C	0.15	−0.044	0.007	7.48×10^{-10}	0.12	0.95	−0.166	0.054	1.04×10^{-3}
let-7e-5p	rs2198171	19	52174483	G	0.27	−0.089	0.014	3.10×10^{-10}	0.25	0.97	−0.124	0.043	1.83×10^{-3}
miR-197-3p	rs7355073	1	110129740	T	0.16	−0.078	0.011	1.23×10^{-12}	0.19	1.00	−0.118	0.041	2.10×10^{-3}
miR-26b-5p	rs12623740	2	219665715	A	0.49	−0.060	0.007	3.37×10^{-18}	0.51	0.99	−0.138	0.051	3.24×10^{-3}
miR-152-3p	rs9910516	17	46183160	A	0.23	0.093	0.016	1.52×10^{-08}	0.27	0.95	0.089	0.033	3.44×10^{-3}
miR-27b-3p	rs10993381	9	97639463	T	0.07	0.170	0.016	2.00×10^{-24}	0.06	0.99	0.148	0.055	3.86×10^{-3}
miR-182-5p	rs2693738	7	129431977	G	0.32	0.115	0.020	2.36×10^{-08}	0.37	0.82	0.166	0.063	4.30×10^{-3}
miR-181a-3p	rs1434282	1	199010721	C	0.27	0.211	0.022	9.03×10^{-21}	0.26	0.98	0.122	0.048	5.57×10^{-3}
miR-181a-5p	rs12125200	1	198992043	A	0.27	0.340	0.013	1.13×10^{-111}	0.24	0.96	0.124	0.049	5.79×10^{-3}
miR-584-5p	rs4147470	5	148528107	T	0.49	−0.131	0.014	7.71×10^{-20}	0.51	1.00	−0.081	0.032	6.15×10^{-3}
miR-26b-5p	rs833083	2	219336959	T	0.41	−0.076	0.006	3.96×10^{-30}	0.43	0.81	−0.137	0.057	7.96×10^{-3}
miR-181a-5p	rs878254	1	199257141	A	0.48	−0.122	0.015	3.54×10^{-15}	0.49	0.90	−0.104	0.045	0.010
miR-181a-5p	rs2360961	1	199000277	C	0.40	−0.151	0.016	4.39×10^{-20}	0.40	0.94	−0.095	0.043	0.014
miR-30d-5p	rs13282464	8	135707922	T	0.15	0.092	0.007	2.02×10^{-33}	0.17	1.00	0.047	0.023	0.020
miR-4433b-5p	rs6740438	2	64528086	C	0.13	0.163	0.029	1.78×10^{-08}	0.15	0.98	0.083	0.041	0.022
miR-30d-5p	rs13268530	8	135727196	T	0.15	0.095	0.007	1.68×10^{-35}	0.17	0.99	0.045	0.023	0.024
miR-21-5p	rs2665392	17	57809453	A	0.16	0.059	0.011	3.59×10^{-08}	0.16	0.88	0.078	0.041	0.027
miR-4433b-5p	rs35503140	2	64539015	C	0.21	−0.130	0.022	9.86×10^{-09}	0.19	0.95	−0.071	0.037	0.029
miR-584-5p	rs9325124	5	148248818	A	0.39	−0.085	0.015	7.62×10^{-09}	0.45	1.00	−0.056	0.031	0.036
miR-181a-5p	rs3861924	1	199121330	A	0.18	0.137	0.020	2.06×10^{-11}	0.20	0.96	0.097	0.054	0.037
miR-1908-5p	rs174561	11	61582708	C	0.30	0.151	0.012	4.76×10^{-31}	0.26	1.00	0.052	0.030	0.040
miR-151a-3p	rs11167012	8	141968408	A	0.42	0.059	0.006	3.79×10^{-24}	0.40	1.00	0.061	0.036	0.045
miR-139-3p	rs10898849	11	72269302	T	0.25	0.124	0.022	3.30×10^{-08}	0.27	1.00	0.054	0.032	0.046
let-7i-5p	rs6581454	12	62934442	G	0.47	0.039	0.006	3.04×10^{-11}	0.44	0.99	0.034	0.021	0.049
Trans associations													
miR-222-3p	rs11070216	15	39817245	T	0.19	−0.067	0.012	4.87×10^{-08}	0.19	0.97	−0.198	0.051	5.06×10^{-5}
miR-222-3p	rs970280	15	39864403	G	0.32	−0.064	0.010	8.79×10^{-10}	0.32	0.94	−0.113	0.042	3.57×10^{-3}
miR-143-3p	rs4734879	8	106583124	G	0.28	0.239	0.031	2.88×10^{-14}	0.24	0.96	0.098	0.038	5.60×10^{-3}
miR-1-3p	rs11906462	20	61158952	T	0.20	0.310	0.033	6.28×10^{-20}	0.23	0.42	0.262	0.116	0.012
miR-320a	rs1443651	2	68569316	G	0.45	−0.036	0.006	7.12×10^{-10}	0.44	1.00	−0.053	0.028	0.029
miR-16-5p	rs137214	22	35288857	T	0.28	0.041	0.007	1.76×10^{-08}	0.29	0.97	0.088	0.050	0.040
miR-126-3p	rs600038	9	136151806	C	0.21	0.055	0.009	5.95×10^{-09}	0.34	1.00	0.041	0.024	0.041
miR-320c	rs1443651	2	68569316	G	0.45	−0.031	0.005	2.77×10^{-10}	0.44	1.00	−0.066	0.039	0.045

[a] One-sided test *P*-value.

EA, effect allele; EAF, effect allele frequency.

the PDCL3 transcript levels in different tissues according to the GTeX database.[41] Another intronic miSNP located in the *NELL1* gene was associated with hsa-miR-320d levels. The seven other *trans* eQTL are located in intergenic regions.

We sought to *in silico* replicate these miSNP associations using the results from Nikpay *et al.*[31] who scanned for genetic polymorphisms associated with miRNA levels in 710 plasma samples. Unfortunately, as the Nikpay *et al.* study relied on a genotyping array focusing mainly on coding regions and used a very stringent imputation quality criterion ($r^2 > 0.9$), it was not possible to assess all our candidate associations. Only four were testable (hsa-miR-941 × rs140930133, hsa-miR-432-5p × rs201969986, hsa-miR-654-5p × rs11109171, hsa-miR-320c × rs10151482) among which only the association of rs140930133 with hsa-miR-941 levels replicated ($P = 6.3 \times 10^{-11}$).

Conversely, we looked into the MARTHA results to replicate the 223 miSNP associations that were significantly ($P < 5 \times 10^{-8}$) detected in the Nikpay *et al.* study. We were able to test 92 of them among which 37 replicated at the nominal level of $P = 0.05$ in MARTHA (*Table 4*). These involved 29 *cis* and 8 *trans* miSNP associations.

Among these eight *trans* miSNP associations, three deserve to be highlighted. First, plasma levels of hsa-miR-143-3p were influenced by the intronic *ZFPM2* rs4734879, *ZFPM2* being a locus reported to associate with venous thrombosis risk[42] and platelet function.[43] In MARTHA, plasma levels of hsa-miR-143-3p were negatively significantly correlated with BMI ($\rho = -0.24$, $P = 3.6 \times 10^{-4}$) and borderline significant with PAI-1 activity levels ($\rho = -0.21$, $P = 5.3 \times 10^{-3}$) (Supplementary material online, *Table S2*). Second, hsa-miR-126-3p plasma levels were associated with the rs600038 located in the promoter region of the *ABO* gene. This polymorphism is in strong LD with several other *ABO* polymorphisms that are known to associate with VT risk, including the rs579459 ($r^2 = 0.99$) tagging for the A1 ABO blood group. In MARTHA, plasma levels of hsa-miR-126-3p were strongly and positively correlated ($\rho \sim 0.20$) with red cells ($P = 1.73 \times 10^{-5}$), lymphocytes ($P = 2.5 \times 10^{-4}$), platelets ($P = 5.9 \times 10^{-4}$), and polynuclear ($P = 6.0 \times 10^{-4}$) (Supplementary material online, *Table S2*). Third, polymorphisms (rs970280, rs11070216) in the promoter region of the *THBS1* gene were found associated with plasma levels of hsa-miR-222-3p. This miRNA has been previously reported to associate with the risk of VT recurrence[13] and has a suggestive association ($P = 0.049$) in our study (Supplementary material online, *Table S3*), where it positively correlated with antithrombin levels ($\rho = 0.21$, $P = 8.8 \times 10^{-4}$) (Supplementary material online, *Table S2*). THBS1 encodes Thrombospondin-1 and is known to be involved in angiogenesis and platelet aggregation.[44,45]

Finally, we performed a random-effect meta-analysis of both datasets in order to discover additional miSNPs. At the 5×10^{-8} statistical threshold, we identified seven new *cis* and five new *trans* miSNP associations (*Table 5*). None of these miSNP associations appeared to involve loci with documented link with thrombosis related traits.

Discussion and conclusion

In this study, we reported the largest investigation to date of miRNA plasma profiling in a cohort of VT patients. Capitalizing on the application of a next-generation sequencing technology, known to be more efficient and sensitive to detect and quantify miRNAs compared with microarray or RT-qPCR techniques, we were able to detect 162 highly expressed miRNAs. These miRNAs were then tested for association with several VT-related phenotypes including 38 haematological traits and VT recurrence. In order to deal with the correlation between miRNA levels and reduce the multiple testing burden associated with the number of tested miRNAs, we deployed an original BN analysis aimed at identifying miRNAs that could serve as more powerful biomarkers for the investigated traits. In addition, as our studied VT patients had been previously typed for genome-wide genotypes, we were able to perform GWAS on each of the 162 miRNAs, and combined our results with some previously obtained in disease-free individuals in order to identify novel associations of common SNPs with plasma miRNA levels.

Several conclusions could be derived from this work. First, we did not identify any miRNA that significantly associated with the risk of VT recurrence. In our study, the miRNA that discriminated the most between patients with or without recurrence, but also between DVT vs. PE patients, was the hsa-miR-370-3p. Several works have already reported the involvement of has-miR-370-3p in lipids metabolism[46-49] and one of the most robust target gene for hsa-miR-370-3p is *CPT1A*[50] whose role in lipid metabolism is also very documented.[51-53] Hsa-miR-370-3p is also predicted to target drug-metabolism genes, such as *CYP2D6* and *VKORC1L1*,[50] that are related to the warfarin anticoagulant pharmacotherapy. Aside this miRNA, we observed a trend of association with VT recurrence for the hsa-mir-27b-3p and hsa-miR-222-3p that had been previously identified in Wang *et al.*[13] but these associations ($P = 0.016$ and $P = 0.0495$, respectively) did not survive any multiple testing correction (Supplementary material online, *Table S3*). Larger studies would be mandatory to confirm these observations and increase our chance to identify other miRNAs associated with the risk of recurrence in VT patients. Second, we observed several significant associations of miRNAs with haematological traits that deserve further replication in independent studies. One can highlight the significant correlation between haematocrit levels and plasma levels of hsa-miR-199b-3p, a miRNA that has been reported to be associated with VT risk.[12] Third, our miR-QTL study identified about 25 significant ($P < 5 \times 10^{-8}$) associations of SNPs with plasma miRNA levels, of which, to the best of our knowledge, 21 have never been reported, including a dozen of *trans* associations. These associations could help deciphering the genomic architecture of complex diseases where miRNAs are involved. For example, plasma levels of hsa-miR-143-3p were found to be associated with the rs4734879 mapping to *ZFPM2*, a gene known to associate with platelet function[43] and VT risk.[42] We also observed a strong association of rs12473206 with plasma levels of hsa-miR-4433b-3p, a miRNA whose serum levels have recently shown to be associated with stroke.[54] The impact of this SNP on stroke risk deserves to be further and deeply investigated. The results of our GWAS on miRNA levels were combined with those obtained by Nikpay *et al.*[31] and freely available at https:// zenodo. org/ . However, only SNPs with imputation quality greater than 0.90 are available at this resource, which has hampered our ability to replicate some of the main associations observed in the MARTHA miRNA study. To facilitate future studies aimed at disentangling the genetic regulation of miRNAs, the results of the 162 GWAS performed on miRNA levels in MARTHA will be available for download at https://zenodo.org/.

Altogether, this study produced a rich source of information relating to plasma miRNAs and biological/clinical traits associated with VT that could be of great use to generate and/or validate new hypothesis.

Supplementary material

Supplementary material is available at *European Heart Journal-Supplement* online.

Table 5 Significant ($P < 5 \times 10^{-8}$) associations of miSNP with miRNA plasma levels derived from the MARTHA miRNA and Nikpay *et al.*[31] meta-analysis

					MARTHA					Nikpay				Combined			
miRNA	chr	Position (bp)	SNP	EA	EAF	r^2	β	SE	*P*	EAF	β	SE	*P*	P^{a}	β	SE	P^{b}
Cis associations																	
miR-181b-5p	1	199257141	rs878254	A	0.485	0.90	−0.054	0.032	0.0916	0.480	−0.071	0.013	$1.64\ 10^{-7}$	0.61	−0.069	0.012	$3.18\ 10^{-8}$
miR-148a-3p	7	25991977	rs9639523	T	0.375	0.87	−0.081	0.034	0.0191	0.344	−0.072	0.013	$2.03\ 10^{-7}$	0.80	−0.073	0.013	$8.41\ 10^{-9}$
let-7a-5p	9	96916230	rs10512230	T	0.287	1.00	0.040	0.031	0.1934	0.315	0.026	0.004	$6.49\ 10^{-8}$	0.67	0.027	0.005	$2.19\ 10^{-8}$
let-7d-5p	9	97229465	rs4497033	T	0.492	0.99	−0.061	0.036	0.0895	0.463	−0.028	0.005	$1.50\ 10^{-7}$	0.36	−0.029	0.005	$3.85\ 10^{-8}$
miR-2110	10	115933905	rs17091403	T	0.091	1.00	−0.141	0.043	$1.13\ 10^{-3}$	0.074	−0.103	0.023	$9.90\ 10^{-6}$	0.44	−0.112	0.020	$4.34\ 10^{-8}$
miR-342-3p	14	100256449	rs8011282	C	0.474	0.99	0.095	0.030	$1.39\ 10^{-3}$	0.487	0.067	0.014	$5.65\ 10^{-6}$	0.41	0.073	0.013	$3.68\ 10^{-8}$
miR-99b-5p	19	52160843	rs11084100	C	0.392	1.00	−0.067	0.024	$5.17\ 10^{-3}$	0.419	−0.065	0.012	$1.12\ 10^{-7}$	0.94	−0.066	0.011	$1.50\ 10^{-9}$
Trans associations																	
miR-215-5p	2	171402733	rs724806	C	0.252	0.97	0.091	0.057	0.1123	0.326	0.143	0.027	$1.44\ 10^{-7}$	0.40	0.134	0.024	$4.09\ 10^{-8}$
miR-10b-5p	7	13236107	rs6948643	G	0.264	1.00	−0.071	0.040	0.0766	0.285	−0.09	0.017	$2.84\ 10^{-7}$	0.66	−0.087	0.016	$4.62\ 10^{-8}$
let-7d-3p	11	2611449	rs1024164	A	0.133	0.87	−0.083	0.034	0.0147	0.092	−0.065	0.013	$7.78\ 10^{-7}$	0.63	−0.068	0.012	$3.18\ 10^{-8}$
miR-378a-3p	11	133763476	rs10894759	A	0.317	0.99	0.066	0.028	0.0206	0.296	0.059	0.011	$7.86\ 10^{-7}$	0.82	0.060	0.011	$3.58\ 10^{-8}$
miR-7-5p	15	41614621	rs7163989	G	0.293	0.99	−0.112	0.041	$6.68\ 10^{-3}$	0.278	−0.089	0.016	$1.48\ 10^{-7}$	0.61	−0.093	0.016	$2.70\ 10^{-9}$

EAF, estimated allele frequency; r^2, imputation quality criterion; β, allele effect.

[a] *P*-value of the test for heterogeneity between the MARTHA and Nikpay studies.

[b] *P*-value of the combined effect obtained through a random-effect meta-analysis of the results of both studies.

Funding

F.T., G.M., and M.G. were financially supported by the GENMED Laboratory of Excellence on Medical Genomics (ANR-10-LABX-0013). D.A.T. was financially supported by the «EPIDEMIOM-VTE» Senior Chair from the Initiative of Excellence of the University of Bordeaux. MiRNA sequencing in the MARTHA study was performed on the iGenSeq platform (ICM Institute, Paris) and supported by a grant from the European Society of Cardiology for Medical Research Innovation. Bioinformatics and statistical analyses benefit from the CBiB computing centre of the University of Bordeaux. This paper was published as part of a supplement supported by an educational grant from Boehringer Ingelheim.

Conflict of interest: none declared.

References

1. Goldhaber SZ. Venous thromboembolism: epidemiology and magnitude of the problem. *Best Pract Res Clin Haematol* 2012;**25**:235–242.
2. Galanaud J-P, Monreal M, Kahn SR. Epidemiology of the post-thrombotic syndrome. *Thromb Res* 2018;**164**:100–109.
3. White RH. The epidemiology of venous thromboembolism. *Circulation* 2003;**107**:41–48.
4. Prandoni P, Bernardi E, Marchiori A, Lensing AWA, Prins MH, Villalta S, Bagatella P, Sartor D, Piccioli A, Simioni P, Pagnan A, Girolami A. The long term clinical course of acute deep vein thrombosis of the arm: prospective cohort study. *BMJ* 2004;**329**:484–485.
5. Kearon C, Parpia S, Spencer FA, Schulman S, Stevens SM, Shah V, Bauer KA, Douketis JD, Lentz SR, Kessler CM, Connors JM, Ginsberg JS, Spadafora L, Julian JA. Long-term risk of recurrence in patients with a first unprovoked venous thromboembolism managed according to d-dimer results; a cohort study. *J Thromb Haemost* 2019; **17**:1144–1152.
6. Bartel DP. Metazoan microRNAs. *Cell* 2018;**173**:20–51.
7. McManus DD, Freedman JE. MicroRNAs in platelet function and cardiovascular disease. *Nat Rev Cardiol* 2015;**12**:711–717.
8. Marchand A, Proust C, Morange P-E, Lompré A-M, Trégouët D-A. miR-421 and miR-30c inhibit SERPINE 1 gene expression in human endothelial cells. *PLoS One* 2012;**7**:e44532.
9. Arroyo AB, Los Reyes-García AM, de Teruel-Montoya R, Vicente V, González-Conejero R, Martínez C. microRNAs in the haemostatic system: more than witnesses of thromboembolic diseases? *Thromb Res* 2018;**166**:1–9.
10. Vossen CY, Hylckama Vlieg A, van Teruel-Montoya R, Salloum-Asfar S, Haan H, de Corral J, Reitsma P, Koeleman BPC, Martínez C. Identification of coagulation gene 3’UTR variants that are potentially regulated by microRNAs. *Br J Haematol* 2017;**177**:782–790.
11. Sennblad B, Basu S, Mazur J, Suchon P, Martinez-Perez A, Hylckama Vlieg A, van Truong V, Li Y, Gådin JR, Tang W, Grossman V, Haan HG, de Handin N, Silveira A, Souto JC, Franco-Cereceda A, Morange P-E, Gagnon F, Soria JM, Eriksson P, Hamsten A, Maegdefessel L, Rosendaal FR, Wild P, Folsom AR, Trégouët D-A, Sabater-Lleal M. Genome-wide association study with additional genetic and post-transcriptional analyses reveals novel regulators of plasma factor XI levels. *Hum Mol Genet* 2017;**26**:637–649.
12. Starikova I, Jamaly S, Sorrentino A, Blondal T, Latysheva N, Sovershaev M, Hansen J-B. Differential expression of plasma miRNAs in patients with unprovoked venous thromboembolism and healthy control individuals. *Thromb Res* 2015;**136**:566–572.
13. Wang X, Sundquist K, Svensson PJ, Rastkhani H, Palmér K, Memon AA, Sundquist J, Zöller B. Association of recurrent venous thromboembolism and circulating microRNAs. *Clin Epigenetics* 2019;**11**:28.
14. Oudot-Mellakh T, Cohen W, Germain M, Saut N, Kallel C, Zelenika D, Lathrop M, Trégouët D-A, Morange P-E. Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol* 2012;**157**:230–239.
15. Germain M, Saut N, Oudot-Mellakh T, Letenneur L, Dupuy A-M, Bertrand M, Alessi M-C, Lambert J-C, Zelenika D, Emmerich J, Tiret L, Cambien F, Lathrop M, Amouyel P, Morange P-E, Trégouët D-A. Caution in interpreting results from imputation analysis when linkage disequilibrium extends over a large distance: a case study on venous thrombosis. *PLoS One* 2012;**7**:e38538.
16. Germain M, Chasman DI, de Haan H, Tang W, Lindström S, Weng L-C, de Andrade M, de Visser MCH, Wiggins KL, Suchon P, Saut N, Smadja DM, Le Gal G, van Hylckama Vlieg A, Di Narzo A, Hao K, Nelson CP, Rocanin-Arjo A, Folkersen L, Monajemi R, Rose LM, Brody JA, Slagboom E, Aïssi D, Gagnon F, Deleuze J-F, Deloukas P, Tzourio C, Dartigues J-F, Berr C, Taylor KD, Civelek M, Eriksson P, Psaty BM, Houwing-Duitermaat J, Goodall AH, Cambien F, Kraft P, Amouyel P, Samani NJ, Basu S, Ridker PM, Rosendaal FR, Kabrhel C, Folsom AR, Heit J, Reitsma PH, Trégouët D-A, Smith NL, Morange P-E. Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet* 2015;**96**:532–542.
17. Thibord F, Perret C, Roux M, Suchon P, Germain M, Deleuze J-F, Morange P-E, Trégouët D-A; GENMED Consortium. OPTIMIR, a novel algorithm for integrating available genome-wide genotype data into miRNA sequence alignment analysis. *RNA* 2019;**25**:657–668.
18. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 2014; **42**:D68–D73.
19. Ameres SL, Zamore PD. Diversifying microRNA sequence and function. *Nat Rev Mol Cell Biol* 2013;**14**:475–488.
20. Kirschner MB, Edelman JJB, Kao S-H, Vallely MP, Van Zandwijk N, Reid G. The impact of hemolysis on cell-free microRNA. *Biomarkers. Front Genet* 2013;**4**:94.
21. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**:550.
22. Ramachandran P, Sánchez-Taltavull D, Perkins TJ. Uncovering robust patterns of microRNA co-expression across cancers using Bayesian Relevance Networks. *PLoS One* 2017;**12**:e0183103.
23. Töpner K, Rosa GJM, Gianola D, Schön C-C. Bayesian networks illustrate genomic and residual trait connections in maize (Zea mays L.). *G3 GenesGenomesGenetics* 2017;**7**:2779–2789.
24. Scutari M. Learning Bayesian Networks with the bnlearn R Package. *J Stat Softw* 2010;**35**:1–22.
25. Florijn BW, Bijkerk R, van der Veer EP, van Zonneveld AJ. Gender and cardiovascular disease: are sex-biased microRNA networks a driving force behind heart failure with preserved ejection fraction in women? *Cardiovasc Res* 2018;**114**:210–225.
26. Huan T, Chen G, Liu C, Bhattacharya A, Rong J, Chen BH, Seshadri S, Tanriverdi K, Freedman JE, Larson MG, Murabito JM, Levy D. Age-associated microRNA expression in human peripheral blood is associated with all-cause mortality and age-related traits. *Aging Cell* 2018;**17**:e12687.
27. Chen X, Liang H, Guan D, Wang C, Hu X, Cui L, Chen S, Zhang C, Zhang J, Zen K, Zhang C-Y. A combination of Let-7d, Let-7g and Let-7i serves as a stable reference for normalization of serum microRNAs. *PLoS One* 2013;**8**:e79652.
28. Tsai W-Y, Jewell NP, Wang M-C. A note on the product-limit estimator under right censoring and left truncation. *Biometrika* 1987;**74**:883–886.
29. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 2005;**95**:221–227.
30. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010;**34**:816–834.
31. Nikpay M, Beehler K, Valsesia A, Hager J, Harper M-E, Dent R, McPherson R. Genome-wide identification of circulating-miRNA expression quantitative trait loci reveals the role of several miRNAs in the regulation of Cardiometabolic phenotypes. *Cardiovasc Res* 2019; **115**:1629–1645.
32. Mägi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 2010;**11**:288.
33. Rubio M, Bustamante M, Hernandez-Ferrer C, Fernandez-Orth D, Pantano L, Sarria Y, Piqué-Borras M, Vellve K, Agramunt S, Carreras R, Estivill X, Gonzalez JR, Mayor A. Circulating miRNAs, isomiRs and

small RNA clusters in human plasma and breast milk. *PLoS One* 2018; **13**:e0193527.

34. Braekkan SK, Mathiesen EB, Njølstad I, Wilsgaard T, Hansen J-B. Hematocrit and risk of venous thromboembolism in a general population. The Tromso study. *Haematologica* 2010;**95**:270–275.
35. Rezende SM, Lijfering WM, Rosendaal FR, Cannegieter SC. Hematologic variables and venous thrombosis: red cell distribution width and blood monocyte count are associated with an increased risk. *Haematologica* 2014;**99**:194–200.
36. Huan T, Rong J, Liu C, Zhang X, Tanriverdi K, Joehanes R, Chen BH, Murabito JM, Yao C, Courchesne P, Munson PJ, O'Donnell CJ, Cox N, Johnson AD, Larson MG, Levy D, Freedman JE. Genome-wide identification of microRNA expression quantitative trait loci. *Nat Commun* 2015;**6**:6601.
37. Civelek M, Hagopian R, Pan C, Che N, Yang W, Kayne PS, Saleem NK, Cederberg H, Kuusisto J, Gargalovic PS, Kirchgessner TG, Laakso M, Lusis AJ. Genetic regulation of human adipose microRNA expression and its consequences for metabolic traits. *Hum Mol Genet* 2013; **22**:3023–3037.
38. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, Oliver-Williams C, Kamat MA, Prins BP, Wilcox SK, Zimmerman ES, Chi A, Bansal N, Spain SL, Wood AM, Morrell NW, Bradley JR, Janjic N, Roberts DJ, Ouwehand WH, Todd JA, Soranzo N, Suhre K, Paul DS, Fox CS, Plenge RM, Danesh J, Runz H, Butterworth AS. Genomic atlas of the human plasma proteome. *Nature* 2018;**558**:73–79.
39. Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, Ingle CE, Dunning M, Flicek P, Koller D, Montgomery S, Tavaré S, Deloukas P, Dermitzakis ET. Population genomics of human gene expression. *Nat Genet* 2007;**39**:1217–1224.
40. Wang Y-P, Li K-B. Correlation of expression profiles between microRNAs and mRNA targets using NCI-60 data. *BMC Genomics* 2009;**10**:218.
41. GTEx Consortium. The genotype-tissue expression (GTEx) project. *Nat Genet* 2013;**45**:580–585.
42. Klarin D, Emdin CA, Natarajan P, Conrad MF, Kathiresan S. Genetic analysis of venous thromboembolism in UK Biobank identifies the ZFPM2 locus and implicates obesity as a causal risk factor. *Circ Cardiovasc Genet* 2017;**10**. pii: e001643.
43. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, Lambourne JJ, Sivapalaratnam S, Downes K, Kundu K, Bomba L, Berentsen K, Bradley JR, Daugherty LC, Delaneau O, Freson K, Garner SF, Grassi L, Guerrero J, Haimel M, Janssen-Megens EM, Kaan A, Kamat M, Kim B, Mandoli A, Marchini J, Martens JHA, Meacham S, Megy K, O'Connell J, Petersen R, Sharifi N, Sheard SM, Staley JR, Tuna S, van der Ent M, Walter K, Wang S-Y, Wheeler E, Wilder SP, Iotchkova V, Moore C, Sambrook J, Stunnenberg HG, Di Angelantonio E, Kaptoge S, Kuijpers TW, Carrillo-de-Santa-Pau E, Juan D, Rico D, Valencia A, Chen L, Ge B, Vasquez L, Kwan T, Garrido-Martín D, Watt S, Yang Y, Guigo R, Beck S, Paul DS, Pastinen T, Bujold D, Bourque G, Frontini M, Danesh J, Roberts DJ, Ouwehand WH, Butterworth AS, Soranzo N. The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* 2016;**167**:1415–1429.e19.
44. Lawler PR, Lawler J. Molecular basis for the regulation of angiogenesis by thrombospondin-1 and -2. *Cold Spring Harb Perspect Med* 2012;**2**:a006627.
45. Trumel C, Plantavid M, Lévy-Tolédano S, Ragab A, Caen JP, Aguado E, Malissen B, Payrastre B. Platelet aggregation induced by the C-terminal peptide of thrombospondin-1 requires the docking protein LAT but is largely independent of alphaIIb/beta3. *J Thromb Haemost* 2003;**1**:320–329.
46. Iliopoulos D, Drosatos K, Hiyama Y, Goldberg IJ, Zannis VI. MicroRNA-370 controls the expression of microRNA-122 and Cpt1alpha and affects lipid metabolism. *J Lipid Res* 2010;**51**:1513–1523.
47. Gao W, He H-W, Wang Z-M, Zhao H, Lian X-Q, Wang Y-S, Zhu J, Yan J-J, Zhang D-G, Yang Z-J, Wang L-S. Plasma levels of lipometabolism-related miR-122 and miR-370 are increased in patients with hyperlipidemia and associated with coronary artery disease. *Lipids Health Dis* 2012;**11**:55.
48. Benatti RO, Melo AM, Borges FO, Ignacio-Souza LM, Simino L. A P, Milanski M, Velloso LA, Torsoni MA, Torsoni AS. Maternal high-fat diet consumption modulates hepatic lipid metabolism and microRNA-122 (miR-122) and microRNA-370 (miR-370) expression in offspring. *Br J Nutr* 2014;**111**:2112–2122.
49. Tian D, Sha Y, Lu J-M, Du X-J. MiR-370 inhibits vascular inflammation and oxidative stress triggered by oxidized low-density lipoprotein through targeting TLR4. *J Cell Biochem* 2018; **119**:6231–6237.
50. Chou C-H, Shrestha S, Yang C-D, Chang N-W, Lin Y-L, Liao K-W, Huang W-C, Sun T-H, Tu S-J, Lee W-H, Chiew M-Y, Tai C-S, Wei T-Y, Tsai T-R, Huang H-T, Wang C-Y, Wu H-Y, Ho S-Y, Chen P-R, Chuang C-H, Hsieh P-J, Wu Y-S, Chen W-L, Li M-J, Wu Y-C, Huang X-Y, Ng FL, Buddhakosai W, Huang P-C, Lan K-C, Huang C-Y, Weng S-L, Cheng Y-N, Liang C, Hsu W-L, Huang H-D. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res* 2018;**46**:D296–D302.
51. Gagnon F, Aïssi D, Carrié A, Morange P-E, Trégouët D-A. Robust validation of methylation levels association at CPT1A locus with lipid plasma levels1. *J Lipid Res* 2014;**55**:1189–1191.
52. Frazier-Wood AC, Aslibekyan S, Absher DM, Hopkins PN, Sha J, Tsai MY, Tiwari HK, Waite LL, Zhi D, Arnett DK. Methylation at CPT1A locus is associated with lipoprotein subfraction profiles. *J Lipid Res* 2014;**55**:1324–1330.
53. Irvin MR, Zhi D, Joehanes R, Mendelson M, Aslibekyan S, Claas SA, Thibeault KS, Patel N, Day K, Jones LW, Liang L, Chen BH, Yao C, Tiwari HK, Ordovas JM, Levy D, Absher D, Arnett DK. Epigenome-wide association study of fasting blood lipids in the Genetics of Lipid-lowering Drugs and Diet Network study. *Circulation* 2014; **130**:565–572.
54. Sonoda T, Matsuzaki J, Yamamoto Y, Sakurai T, Aoki Y, Takizawa S, Niida S, Ochiya T. Serum microRNA-based risk prediction for stroke. *Stroke* 2019;**50**:1510–1518.